

Database Security of the Automated Control Subsystem of Water Conservation Processes

Vitalii Osipa^{1,†}, Lesya Ladieva^{2,†} and Bogdan Korniyenko^{2,*}

¹ Pantheon-Sorbonne University, Place du Panthéon 12, 75231, Paris, France

² National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Beresteyskyi Avenue, 37, 03056, Kyiv, Ukraine

Abstract

The article considers the issues of database security of the automated subsystem of water conservation process management of a modern enterprise. The reason for its complexity is that it must also have connections between different data elements and with ensuring information protection. That is why an important task in database development is the development of a conceptual data model (the general logical structure of the database). An approach is proposed to maintain the viability of this database. The final products of biochemical decomposition of organic matter in such water bodies are mineral salts and water. If the treatment facilities operate effectively, then wastewater after appropriate treatment can be used for technological production purposes. This will significantly reduce the consumption of fresh water, as well as reduce emissions of pollutants into the environment. Automated control systems significantly increase the efficiency of technological purification systems. The quality of management of such processes largely depends on the quality of the automated control database. The components of information support for automated control of water conservation processes of a modern enterprise are presented. The core of information support is the database. The theoretical results of the presented study are specified in the development of a database subschema that functions in the PostgreSQL database management system (DBMS).

Keywords

security, database, data structures, automated control, water conservation

1. Introduction

Water usage after its purification for technological purposes of modern production is advantageous both in terms of environmental protection and economic considerations. If treated wastewater (approximately 90%) is reused for technological purposes, it will be much cheaper for the company than taking fresh water.

Almost always wastewater treatment is a set of methods, among which the functioning of the technological treatment system in terms of using an automated control system, occupies an important place [1,2]. As for the subsystem of automated control of water conservation processes of a modern production, due to the growing volume of data and the complexity of their organization, the effectiveness of its operation depends largely on the reliability of information support. Information support of the subsystem of automated control of water saving processes establishes: the composition of information, determination of information routes, means of information transformation, data structure and conceptual model of the database.

The correctness of the concept on the basis of which the components of information support are formed and the validity of the selection of important features and connections are important for the

Information Technology and Implementation (IT&I-2025), November 20-21, 2025, Kyiv, Ukraine

* Corresponding author.

† These authors contributed equally.

✉ v.osipa@gmail.com (V. Osipa); lrynu@yahoo.com (L. Ladieva); bogdanko@gmx.net (B. Korniyenko)

>ID 0009-0007-2599-8953 (V. Osipa); 0000-0002-1706-0072 (L. Ladieva); 0000-0002-2521-0878 (B. Korniyenko)

 © 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

representativeness of information. The development of information support for the subsystem of automated control of water conservation processes was carried out on the basis of the methodology of programming the life cycle of complex objects [3] (in this case, the technological system of treatment); database development was performed on the basis of the methodology of organization of databases in computer systems, the founder of which is Martin J. [4] and his followers Date K., Krenke D. [5, 6], etc., which has positively affect the practical implementation. According to this methodology, the term schema is used to define a complete table of all types of data elements and record types that will be stored in the database. The term subcircuit defines the description of the data used for a set of individual tasks.

One of the first works where the problem of building a database of automated control of wastewater treatment was considered. It should be noted that this article did not present a logical model of the database. The works [7-9] is also devoted to the problem of developing a similar database. However, in [7], at the third lower level of data presentation, such an important node as "Standard design solutions for typical cleaning processes" was not considered. Note also that both of these publication did not consider the problem of determining the primary and secondary (foreign) keys for the normalized form of data. This fact did not allow to build a complete database model.

2. Information support structure

Structural information support of water-saving processes automated control subsystem for modern production (airport) consists of information support of functional subsystems (ISF) and information support of modernization (design) of technological cleaning system (ISM).

The basis of ISF is a database, the control system of which allows you to develop an interface with applications. Applied software based on economic and mathematical methods and models ensures the functioning of water conservation processes in the optimal mode. ISF consists of system-wide (SWIS) and functionally-autonomous information support, which reflects the specifics of each subsystem of automated control.

System-wide information support consists of:

- classifiers of problems that contain reference data for solving functional problems;
- electronic arrays of these classifiers;
- electronic documents common to all subsystems;
- means of unambiguous description of data used in all subsystems.

Functionally autonomous information support (FAIS) includes:

- information classifiers and corresponding electronic arrays necessary for solving problems in individual subsystems;
- means of unambiguous description of data used in individual subsystems.

Information support for modernization and design of technological treatment system is a set of means of identification of object documentation, stages and stages of work on modernization (design) of technological treatment system, which includes:

- means of unification of names of tasks of modernization and design;
- system of coding of design decisions;
- system of designations and coding of stages and stages of modernization (design) of technological cleaning system;
- coding system of project documentation.

In the future we will consider the functional and autonomous information support of the subsystem of automated control of water conservation processes of a modern enterprise.

The information support of this subsystem includes:

1. normative and reference data (information on maximum permissible concentrations of pollutants, water consumption limit j-m production process, economic assessment of water in the region, unit price of water j-m production process, cost standards provided by scientific and financial plan, etc.);

2. current information for automated control algorithms (wastewater consumption, physicochemical parameters of the contaminant ingredient, reagent characteristics, water

consumption for normal operation of the j -th production process, multiplicity of water dilution by the k -th pollutant, set level of water reuse j - m production process etc.);

3. accounting and archival information for modernization (design) of the technological scheme of wastewater treatment (standard design solutions for typical treatment processes, equipment characteristics, etc.).

As mentioned, the database and management system are the core of information support. The functional use of information support is carried out within the local network, which makes it accessible to all participants in the process. The database administration is responsible for the up-to-dateness of the information, and the functional responsibilities of its employees should be included in the organizational support. Before entering data into the database, their authenticity is checked by special algorithms.

When developing a logical description of the database of the subsystem of automated control of water conservation processes, the sets of data required for control were identified. The data presented at the logical level of the user reflects his view of the database when solving specific tasks. The introduction of this level of data representation significantly reduces the processing time of queries, increases the security of the database and expands the scope of their use.

The scheme of the database of automated control of cleaning systems includes two tree-like structures (Fig. 1).

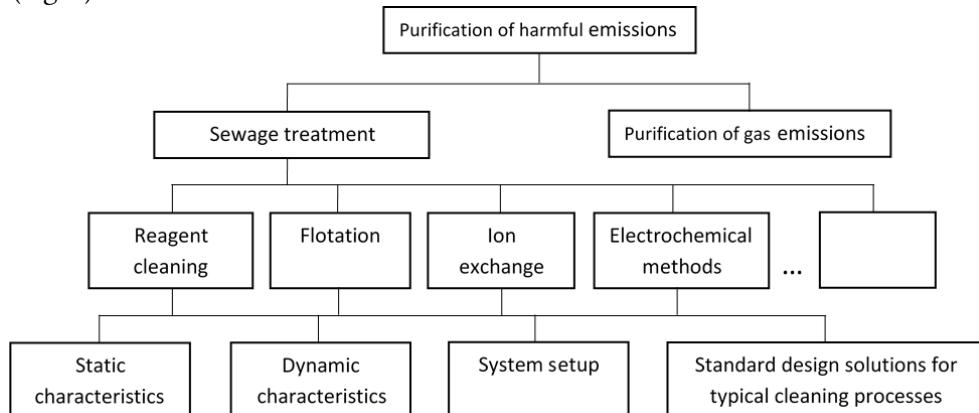


Figure 1: Hierarchical database model of ACS for cleaning of harmful emissions.

The leaf of each tree is an entry that includes the code and name of the treatment (*Wastewater Treatment or Gas Emissions Treatment*) - the top level of the hierarchy. The next level of the hierarchy, subordinate to the first, is represented by records that reflect the methods of cleaning and their codes. The lower level of the hierarchy consists of records for management tasks according to certain cleaning methods.

The minimum available unit of information in the hierarchical model is a node, ie one row at each level. Searching on the basis of a hierarchical data model is carried out in the directions from top to bottom and from left to right. Therefore, to find any node, it is necessary to identify the node that is looking for and those nodes to which it is subordinate. This is a shortcoming of the hierarchical database. In addition, such a database makes it difficult to adjust, enable and delete all subordinate levels, which also complicates the operation of the database. Presentation of data in the form of trees will hinder the development and significant growth of the database. At a certain stage of its growth, it is possible to disrupt the logical representation of data, which will lead to the need for changes in applications. As a result, one of the main advantages of the database will disappear - data independence. This shortcoming is eliminated in a relational database, in which data is presented in the form of two-dimensional tables - one of the most natural ways to present data to the user.

3. Database relations normalization

Since any network structure can be decomposed into a set of tree structures with some redundancy, the hierarchical representation of data can also be reduced to two-dimensional flat files with some

redundancy. Relationships between the data shown in Fig. 2 will be presented in the form of two-dimensional tables. This was done step by step for each connection between the data, using the normalization process. Tables are built in such a way that information about the relationships between data elements is not lost. The relationship key must contain, in addition to the primary, a secondary (foreign) key, with which the tuples of this relationship will be uniquely identified.

For database management systems that do not support static relations of relational relations, the links between them are established only to solve a specific problem and exist for the period of its solution. Therefore, when mapping a hierarchical model to a relational one, all structural connections are not described explicitly, but only a check is made for the possibility of establishing these connections. A prerequisite for establishing a connection between two relational relations is the presence of at least one common attribute, which is the foreign key on which the connection is performed.

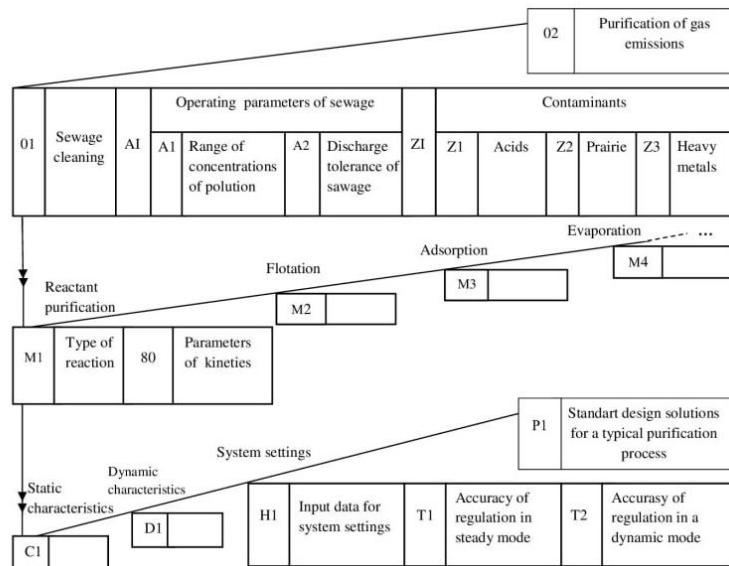


Figure 2: Subschema of ACS database for typical wastewater treatment processes.

Based on the subscheme of the database (fig. 2), the normalized form of the database can be written as follows: initial relation: WASTEWATER COMPONENTS (*wastewater mode parameter code, mode parameters, code of pollutants, type of pollution, purification type code, type of purification*) – (Table 1).

Table 1
Wastewater components

Wastewater mode parameter code Primary Key	MIN mode parameter value	MAX mode parameter value	Code of pollutants Foreign Key	Type of pollution	Purification type code	Type of purification
R1	50	150	Z1	Strong acids and alkalis A mixture of strong and weak acids	01	Wastewater treatment
R1	100	200	Z2	Strong acids and heavy metal ions	01	Wastewater treatment
R1	75	150	Z3	Cyanides and hexavalent chromium	01	Wastewater treatment
R1	80	160	Z4		01	Wastewater treatment

The relationship between PURIFICATION METHODS AND TYPICAL PROCESSES in the first normal form (1NF) is also obtained on the basis of the database subscheme (Table 2).

The relationship will be written as follows: PURIFICATION METHODS AND TYPICAL PROCESSES (code of purification method, name of purification method, code of typical process, name of typical process, code of typical processes parameters, parameters of typical processes, code of pollutant, code of mode parameter).

Table 2

Purification methods and typical processes

Code of purification method	Name of purification method	Code of typical process	Name of typical process	Code of typical processes parameters	Chemical reaction constant	Chemical reaction order	Code of pollutant	Code of mode parameter
M1	Reagent cleaning	60	Neutralization	41	0,125	1	Z1	R1
M1	Reagent cleaning	61	Deposition of heavy metals	42	0,5	2	Z2	R2
M1	Reagent cleaning	62	Disposal	43	0,75	1	Z3	R3
M1	Reagent cleaning	63	Coagulation	44	0,8	1	Z4	R4
M11	Flotation	70	Z11	R11
M21	Ion exchange	80	Z21	R21
M31	Electrochemical methods	90	Z31	R31

Continuing the normalization process, we obtain the ratio of the PURIFICATION METHOD and the ratio of TYPICAL PROCESSES in the third normal form (Table 3).

Table 3

Purification method 3NF

Code of purification method Primary Key	Name of purification method	Code of pollutants Foreign Key	Code of mode parameter
M1	Reagent cleaning	Z1	R1
M2	Flotation	Z20	R20
M3	Ion exchange	Z30	R30
M4	Electrochemical methods	Z40	R40
...	...		

The ratio TYPICAL PROCESSES will first be written in the second normal form (Table 4).

Table 4

Typical processes 2NF

Code of typical process	Name of typical process	Code of typical processes parameters	Chemical reaction constant	Chemical reaction order
60	Neutralization	41	0,125	1
61	Deposition of heavy metals	42	0,5	2
62	Disposal	43	0,75	1
63	Coagulation	44	0,8	1
...

Finally, the ratio TYPICAL PROCESSES in the third normal form will be written (Table 5).

Table 5
Typical processes 3NF

Code of typical process Primary Key	Name of typical process	Code of typical processes parameters	Code of pollutant Foreign Key	Code of purification method	Code of mode parameter
60	Neutralization	41	Z1	M1	R1
61	Deposition of heavy metals	42	Z2	M2	R2
62	Disposal	43	Z3	M3	R3
63	Coagulation	44	Z4	M4	R4
...

The TYPICAL PROCESS PARAMETERS relationship will be written in the third normal form as follows (Table 6).

Table 6
Typical process parameters 3NF

Code of typical processes parameters Primary Key	Chemical reaction constant	Chemical reaction order	Code of pollutant Foreign Key	Code of purification method
41	0,125	1	Z1	M1
42	0,5	2	Z2	M2
43	0,75	1	Z3	M3
44	0,8	1	Z4	M4
...

The generated ratio POLLUTANTS AND THEIR PARAMETERS in the first normal form based on the database subscheme (fig. 2) will be written as follows: POLLUTANTS AND THEIR PARAMETERS (*code of mode parameters, code of pollutant, name of pollutant, physico-chemical parameters of pollutant, pollutant sensor value*) (Table 7).

Table 7
Pollutants and their parameters

Code of mode parameter Primary Key	Code of pollutant Foreign Key	Name of pollutant	Molar mass	Dissociation constant	Density of aqueous solution	Pollutant sensor value [g/l]
R1	Z1	Sulfuric acid	98	$1 \cdot 10^3$	1.2	5
R2	Z2	Strong and weak acids mixture	-	-	1.1	4
R10	Z10	Acetic acid	60	$1,75 \cdot 10^{-5}$	1	3
...
R20	Z20	Iron sulfate Cu^{3+}	152	-	-	10
...

Continuing the process of normalization, the ratio of POLLUTANTS and the ratio of PARAMETERS OF POLLUTANTS in the third normal form were obtained (Table 8-9).

Table 8
Pollutants 3NF

Code of mode parameter Primary Key	Code of pollutant Foreign Key	Name of pollutant
R1	Z1	Sulfuric acid
R2	Z2	Strong and weak acids mixture
R10	Z5	Acetic acid
...
R20	Z10	Iron sulfate Cu ³⁺
...

Table 9
Parameters of pollutants 3NF

Code of mode parameter Primary Key	Code of pollutant Foreign Key	Molar mass	Dissociation constant	Density of aqueous solution	Pollutant sensor value [g/l]
R1	Z1	98	1*10 ³	1,2	5
R2	Z2	-	-	1,1	4
...
R10	Z10	60	1,75*10 ⁻⁵	1	3
R20	Z20	152	-	-	10
...

The generated relation REAGENTS AND THEIR PARAMETERS in the first normal form on the basis of the database subscheme (fig. 2) will be written as follows: REAGENTS AND THEIR PARAMETERS (*reagent type code, reagent name, reagent physicochemical parameters, reagent concentration sensor value, contaminant code, purification method code*) (Table 10).

Table 10
Reagents and their parameters

Reagent type code Primary Key	Reagent name	Molar mass	Density	Reagent concentration sensor value [g/l]	Code of pollutant Foreign Key	Code of purification method
200	Sodium hydroxide	40	1.2	4	Z1	M1
201	Lime suspension	37	1.4	5	Z2	M1
...

Continuing the process of normalization, we obtain the ratio of REAGENTS and the ratio of PARAMETERS OF REAGENTS in the third normal form (Table 11-12).

Table 11
Reagents 3NF

Reagent type code Primary Key	Reagent name	Code of pollutant Foreign Key	Code of purification method
200	Sodium hydroxide	Z1	M1
201	Lime suspension	Z2	M1
...

Table 12
Parameters of reagents 3NF

Reagent type code Primary Key	Molar mass	Density	Reagent concentration sensor value [g/l]	Code of pollutant Foreign Key	Code of purification method
1	2	3	5	6	7
200	40	1.2	4	Z1	M1
101	37	1.4	5	Z2	M1
...

The normalization resulted in a relational database model that combines relationship tables.

4. Database management system selection

The implementation of a logical data model is primarily associated with the selection of a specific database for management tasks (Fig. 3). This task is not easy and many factors need to be considered. Here you need to predict the prospects for production development, which creates a choice in terms of expanding functions and tasks, and studies the software market. There are two problems in the assessment of DBMS. The first is related to the choice of the database from the user's point of view, and the other is purely technical and related to system performance.

At the present time there is no generally accepted analytical method of database management system choice. Therefore, experts in data management field for the solution of this task use such methods:

- methods of simulation;
- experimental studies with using full scale tests;
- heuristic method having such points in the comparative evaluation of characteristics as "yes-no", "present - absent," "good - satisfactory - bad."

While using simulation modeling techniques there can appear some additional problems in evaluation of database simulation model accuracy. Obtaining comparative characteristics of database management system with the help of full-scale test is associated with significant labor costs. Therefore, the choice of database management system for maintaining a database is based on expert estimates (Table 13-14).

Table 13
Expert assessments of promising DBMS

Experts i=1,4	Factors (promising DBMS), j=1,7						
	PostgreSQL	MySQL	MongoDB	SQLite	DB2	MS SQL Server	MariaDB
1	1	3	7	4,5	4,5	6	2
2	1	4	7	3	6	5	2
3	2	1	3	5	6	4	7
4	1	2,5	2,5	5	7	4	6

Table 14
Competence of expert's criteria

Nº Expert	1	2	3	4
Professional experience (years)	6	34	5	5
Qualification (h-index)	4	22	18	5

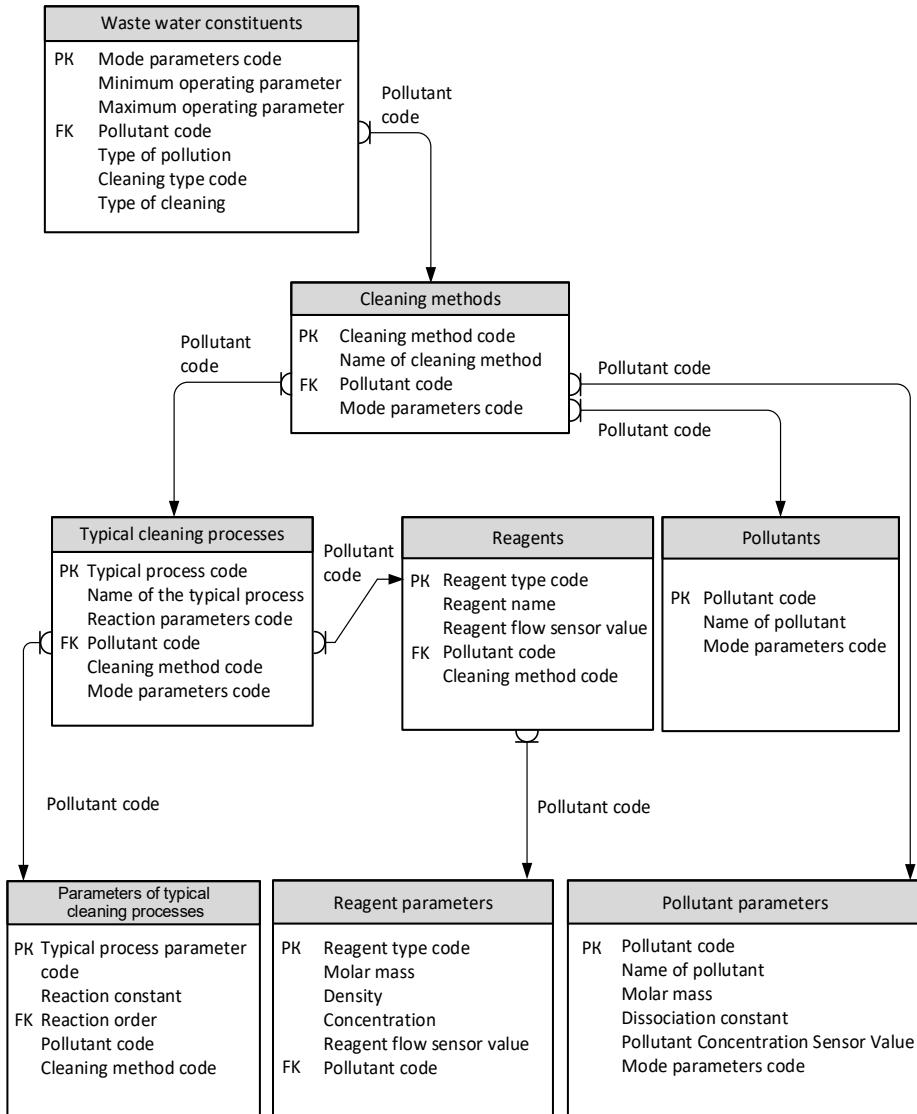


Figure 3: Conceptual model of the database for water-saving processes automated control of modern enterprise.

Taking into account professional experience and qualifications, we assign each rank that 1 is the maximum and 10 is the minimum. Then calculate the coefficients of competence of each of the experts, where 1 - the most competent, 2 - the least competent. These data are translated into ranks on the principle: the longer the work experience or higher qualification, the lower the rank.

Since the matrix of expert assessments contains matching ranks (so-called "linked"), we will first reduce it to the normal form. In a normal matrix, the sum of each row is equal to $K(K+1)/2$. In the (Table 15) shows a normal ranking matrix of experts according to the criteria of competence.

Table 15
Normalized ranking of experts

Nº Expert	1	2	3	4
Professional experience (years)	9,24	1	9,53	9,53
Qualification (h-index)	9,18	1	2,81	8,73
Competence	2	1	1,63	1,9

Table 16
Aggregated results

Experts i=1,4	Factors (promising DBMS), j=1,7							
	PostgreSQL	MySQL	MongoDB	SQLite	DB2	MS SQL Server	MariaDB	Server
1	2	6	14	9	9	12	4	
2	1	4	7	3	6	5	2	
3	3,26	1,63	4,89	8,15	9,78	6,52	11,41	
4	1,9	4,75	4,75	9,5	13,3	7,6	11,4	
Aggregated	8,16	16,38	30,64	29,65	38,08	31,12	28,81	

Therefore, based on the results shown in (Table 16), we can conclude that the PostgreSQL database has the smallest number of ranges. Thus, according to experts, this DBMS is the most suitable time database management system for automated management of water-saving processes. An important feature of the PostgreSQL database is the development of capabilities in its application environment for automated control of languages such as PL/Java, C++, C-compatible modules, Python, PL/pgSQL and others. In addition, the PostgreSQL database has streamlined high-performance transaction and replication mechanisms, as well as an SQL-based query language that allows the client-server LAN to use a client-server architecture. Thus, the PostgreSQL database meets the requirements of efficient operation of the automated control process of water-saving processes for modern production.

5. Protection of information in the subsystem database

Practical work with the presented database involves ensuring three fundamental aspects of security - authentication (who has the right to connect), authorization (what exactly they can do), and protection of the data that is stored and transmitted (Fig.4).

One of the first levels of protection is to verify the client connecting to the database. The PostgreSQL database implements a wide range of authentication methods: trust, password, certificate, as well as integration with external systems - LDAP, GSSAPI/Kerberos, RADIUS, etc. For example, the outdated MD5 method is no longer recommended due to its vulnerability to attacks with "rainbow" tables. Instead, it is recommended to use SCRAM-SHA-256 as a more secure option.

The pg_hba.conf file allows you to flexibly determine: where to connect from (IP addresses or local sockets), which user, to which database, and which authentication method is used.

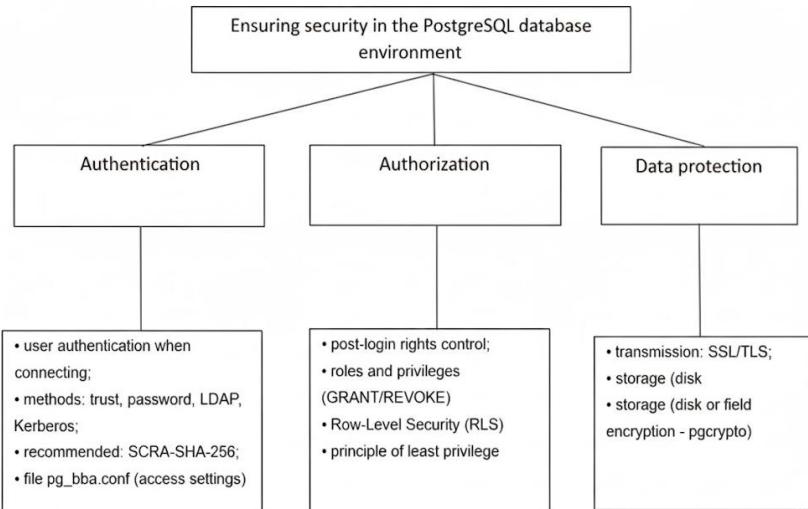


Figure 4: Information protection in the automated control subsystem database

Thus, authentication settings and connection location restrictions reduce the risk of unauthorized access. After confirming the user's authenticity, the next step is to restrict rights: which database objects he can see or change. The PostgreSQL database implements a role and permission model (GRANT/REVOKE) with the principle of "least privilege". In addition, a finer level of control - the so-called Row-Level Security (RLS) - allows you to set policies that control access to individual table rows, depending on the user's role.

The settings of the PUBLIC schema and default privileges also require attention: excessive rights or trust rules in pg_hba.conf can create weak points.

So, authorization is not only about creating roles, but also about regularly auditing rights and reducing rights to the minimum necessary. When it comes to storing and transferring data, the PostgreSQL database has several options, although some of them require external mechanisms.

To protect data during transmission (data in transit), it is recommended to use an SSL/TLS connection, which provides encryption of traffic between the client and the server.

To protect data on disk (data at rest), the PostgreSQL database does not have a full-fledged Transparent Disk Encryption (TDE) feature in the base assembly, so encryption is used at the OS or file system level (for example, LUKS, FileVault) or through an intermediate implementation of TDE based on fork processes. In addition, column or application level encryption can be used (for example, the pgcrypto function), which will allow you to selectively protect sensitive fields. In general, the right strategy is to combine encryption, access restrictions and rights control, which will ensure the integrity, confidentiality and availability of information.

6. Conclusion

Thus, the structure of information support for automated water conservation process management of modern production is proposed, and the composition and structure of data are determined, as well as the choice of a database management system. An approach to ensuring information protection in the database is considered.

The authors have developed a conceptual model of the database as the main component of information support. For each connection between data, normalization was performed, which allowed obtaining a relational data structure in the third normal form, which has significant advantages over hierarchical and network structures.

The client-server architecture of the database for the subsystem of automated water conservation process management will allow not only to store data on the central computer, but also to perform basic data processing operations. The specificity of this architecture is the use of the SQL query

language. The user will not receive files, but only the data selected by him. Protecting information in a database includes measures to ensure its integrity (protection against unauthorized modification), confidentiality (protection against unauthorized disclosure), and availability (ensuring access and recovery from failures). Key techniques include encryption, authentication (including multi-factor authentication), password and privilege management, the use of firewalls and VPNs to create secure connections, and regular backups to recover data in the event of loss.

The presented approaches and the criteria formed on their basis allowed making a choice of a database management system, which allows proceeding to its physical implementation. The choice of a DBMS to support the viability of the database based on the generated indicators is another result of this work. The database will be maintained in the environment of the PostgreSQL relational DBMS, which includes a system of embedded programming languages, the SQL query language, which will allow developing an interface with the application software of the automated control system (ACS) for wastewater treatment processes on a single information basis.

Further development of the database will be associated with the accumulation of information on the statistical and dynamic characteristics of typical treatment processes. Information of the third lower level of the database subscheme will be recorded in the form of sections of the actual database. In the future, an Automated Information System for Managing Water-Saving Processes in Modern Production should function.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

- [1] Bieliatynskyi, A., Osipa, L., Klochan, A., Cheropkin, E. Control algorithm of the airport's water-saving processes. Proceedings of the 21th Conference for Junior Researchers "Science – Future of Lithuania": Transport engineering and management (Vilnius, 4-5 May, Lithuania). 137-141, 2018. URL: <http://jmk.transportas.vgtu.lt/index.php/tran2017/tran2018/paper/viewFile/149/163>
- [2] Osipa R. Statics Simulation of the Buffer Wastewater Neutralization Process. International Journal of Engineering and Manufacturing, Vol. 9, No. 4, 1-14, 2019. doi: 10.5815/ijem.2019.04.01
- [3] Timchenko, A. Fundamentals of system design and system analysis of complex objects. Lubid. Kyiv. 2000.
- [4] Martin, J. Computer data-base organization. Prentice-Hall, Inc., Englewood Cliffs. New Jersey, 1977.
- [5] Date C.J. (2003). Introduction to Database systems. 8th edition , CCNA 20-301 Official Cert., 2003.
- [6] Kroenke D., Auer D., Scott V. Databases processing. Pearson Prentice Hall, 15th edition. 2018.
- [7] Zhuchenko, A., Osipa, L., Cheropkin, E. Design Database for an Automated Control System of Typical Wastewater Treatment Processes. International Journal of Engineering and Manufacturing, Vol.7, No.4, 36-50, 2017. DOI: 10.5815/ijem.2017.04.04.
- [8] Korniyenko, B.Y., Ladieva, L.R., Galata, L.P. Mathematical model of heat transfer process of production of granulated fertilizers in fluidized bed. ARPN Journal of Engineering and Applied Sciences, 16 (20), 2126-2131, 2021.
- [9] Ladieva, L., Kozanevych, Z., Klusta, T., Korniyenko, B. System of control of the process of alkylation of benzene with peripene in the liquid phase. ATIT 2020 - Proceedings: 2020 2nd IEEE International Conference on Advanced Trends in Information Theory, № 9349330, 311-314, 2020.