REPUBLIC OF AZERBAIJAN

On the rights of the manuscript

ABSTRACT

of the dissertation for the degree of Doctor of Philosophy

METHODS OF PROTECTING INFORMATION CONFIDENTIALITY IN SPECIAL-PURPOSE ELECTRONIC DOCUMENT MANAGEMENT SYSTEMS (ON THE EXAMPLE OF MEDICAL INSTITUTIONS)

Specialty:	3338.01 – System analysis, management and information processing
Field of science:	Technical sciences
Applicant:	**Nargiz Firuz Mammadzada**

Baku – 2026

The dissertation was performed at the Department of "Engineering Mathematics and Artificial Intelligence" of the Azerbaijan Technical University.

Scientific supervisor: Doctor of Technical Sciences, Professor
Vagif Alijavad Gasimov

Official opponents:

Doctor of Technical Sciences, Professor
Alakbar Ali Agha Aliyev

Doctor of Technical Sciences, Professor
Balami Gasim Ismayilov

Doctor of Technical Sciences, Professor
Javanshir Firudin Mammadov

Dissertation council FD 2.48 of Supreme Attestation Commission under the President of the Republic of Azerbaijan operating at Azerbaijan State Oil and Industry University.

Chairman of the Dissertation Council: Corr.-Member of ANAS, Doctor of Technical Sciences, Professor
Rafig Aziz Aliyev

Scientific Secretary of the Dissertation Council: Doctor of Technical Sciences, Associate Professor
Akif Vali Alizade

Chairman of the Scientific Seminar: Doctor of Technical Sciences, Professor
Kamala Rafig Aliyeva

İMZANI TƏSDİQ EDİRƏM
ADNSU-nun Elmi katibi

GENERAL CHARACTERISTICS OF THE WORK

Relevance of the topic and degree of development. In the modern era, the widespread adoption of digital technologies and the implementation of electronic document management systems in government, private, and military sectors have brought information security issues to the forefront. In special-purpose electronic document management systems (SPEDMS), the confidentiality, integrity, and availability of information are of critical importance. The application of various cryptographic methods, security protocols, and organizational measures is essential to ensure information security in these systems. One of these areas is related to the documents generated in medical institutions and the process of their secure storage, processing, and transmission. Currently, there are many methods and techniques applied for the security of electronic documents in SPEDMS. However, the weaknesses of existing technologies, as well as emerging cyber threats, necessitate additional research in the field of information security, the development and implementation of new cryptographic methods and algorithms, and security models. Unauthorized access, modification, or destruction of data can cause serious consequences for government and private structures. In addition, aspects such as digital signatures, authentication of electronic documents, and source identification play an important role in the security of SPEDMS.

In particular, information security issues in medical electronic documents are of special importance. Patients' health data is of a confidential nature, and unauthorized access to it can lead to ethical and legal problems. Violation of the confidentiality of medical documents can not only interfere with the patient's personal life but also damage the trust in healthcare institutions. Furthermore, as a result of security breaches in electronic document management systems in the healthcare sector, serious problems such as incorrect diagnoses, interference with the treatment process, and writing of fraudulent prescriptions may arise. International standards and regulatory requirements exist in this area (ISO/IEC 27001, NIST 800-53, HIPAA, etc.), but research has shown that additional

measures are needed to fully implement these standards and comply with their requirements. At the same time, evolving technologies, computing power, and the discovery of vulnerabilities in existing methods over time require the development of new approaches. In this regard, the development and implementation of effective solutions for ensuring information security in SPEDMS, including document management systems in medical institutions, is one of the current scientific and practical problems.

In this dissertation, cryptographic algorithms, security models and architectures, and information confidentiality methods have been developed for ensuring the confidentiality and security of special-purpose electronic documents, storage, processing, and transmission of medical documents using medical institutions as a case study.

Object of the research special-purpose electronic document management systems, using electronic documents in medical institutions as a case study.

Subject of the research consists of creating methods, algorithms, models, and architectures for ensuring the confidentiality and security of medical documents.

Purpose and objectives of the research: It consists of developing cryptographic algorithms, key exchange protocols, architectural models for distributed database architectures, anonymization algorithms for preserving privacy, and system and resource access technologies to ensure confidentiality in the processes of storing, processing, and transmitting electronic documents in medical institutions.

Research methods: In the research conducted in the presented dissertation, chaos theory, random number theory, matrix algebra, statistical analysis methods, and experimental research have been used.

Theoretical and practical significance of the research.

The theoretical and practical significance of the research is characterized by the fundamental and applied nature of the proposed approaches. Specifically, the non-traditional cryptographic framework formed via the chaotic Brownian motion of molecules and the DFS algorithm, along with the alternative key exchange

protocol based on the decomposition of singular matrices, not only establishes a new mathematical environment for data encryption and the reliable anonymization of sensitive medical data while preserving their utility for analysis, but also ensures their protection in real-world databases and secure transmission across networks. Furthermore, the new management models proposed for transaction synchronization in leaderless distributed systems improve the theoretical resolution of data conflicts while achieving practical resource optimization (replication) and fault tolerance by reducing network load through a transaction table-based architecture.

Main provisions submitted for defense:

- Encryption algorithm for text-format documents;
- Cryptographic encryption algorithm for image-format data;
- Protocol for cryptographic key exchange during data transmission in networks;
- Method and algorithm for anonymization of confidential data for data analysis;
- Synchronization and replication architecture for distributed databases;
- Method for creating secure transmission channels for transferring documents to remote databases;
- Authentication and authorization models for resource access.

Scientific novelty of the research:

- Conceptual principles for confidential information exchange between a distributed architecture model for special-purpose electronic document management systems and specially-structured distributed databases have been proposed.
- A new encryption method based on random number generation using the chaotic Brownian motion model of molecules has been developed for ensuring the confidentiality of text-type data (documents) in special-purpose electronic document management systems.
- A new maze-based encryption algorithm generated by the DFS algorithm has been developed for ensuring the confidentiality

of image (graphic, picture) type data (documents) in special-purpose electronic document management systems.

- A special key exchange protocol based on matrix algebra has been developed for ensuring privacy during data transmission in distributed structure database networks.
- A special shuffling algorithm based on maze generation has been developed for ensuring anonymity when using data stored in the databases of the electronic document management system.
- Principles for the use of transaction tables have been proposed to ensure synchronization and replication in distributed structure databases.
- A dynamic roles and permissions approach has been proposed for managing access to data stored in databases.

Approbation and application. The dissertation was carried out's main results were discussed at the following scientific-technical conferences and forums and published:

- Gasimov V.A., Mammadov J.I., Mammadzada N.F. Stream encryption method based on the chaotic brownıan motion model of molecules // 4th International Conference on Innovative Data Communication Technologies and Application (ICIDCA 2022). 3-4 November, http://icidca.com/2022/. Coimbatore, India. Procedia Computer Science, vol.215, 2022, pp. 577-588.;

- Gasimov, Vagif & Mammadzada, Nargiz & Mammadov, Jabir. New Key Exchange Protocol Based on Matrix Algebras. – 2023, 1-3. 10.1109/PCI60110.2023.10326004;

- Gasimov, Vagif, Mammadzada, Nargiz & Mammadov, Jabir. New key exchange protocol based on matrix algebras // II International Conference on Informatıon Securıty: Problems and Prospects, Baku, Azerbaijan, - 2022;

- Gasimov, Vagif, Mammadzada, Nargiz & Mammadov, Jabir (2022). New symmetric encryption algorithm based on chaotic motion of molecules // II International Conference on Informatıon Securıty: Problems and Prospects, Baku, Azerbaijan, – 2022.

- Gasimov Vagif, Mammadzada Nargiz. Data Shuffling Algorithm for Preserving Privacy and Utility of Analysis, ITTA 2026 (3rd International Conference on Information Technologies and Their Applications), Azerbaijan

Name of the organization where the dissertation was carried out. The dissertation was carried out at the "Engineering Mathematics and Artificial Intelligence" Department of the Azerbaijan Technical University.

Total volume of the dissertation in characters, indicating the volume of individual structural sections. The dissertation consists of an introduction of 8,527 characters, Chapter I of 29,766 characters, Chapter II of 103,406 characters, Chapter III of 61,063 characters, and Chapter IV of 31,341 characters, totaling 259,937 characters overall.

BRIEF CONTENT OF THE WORK

In the Introduction, the relevance of the topic was substantiated, and the purpose, scientific novelties, and practical significance of the research were explained.

In the First Chapter, *"Information security problem in special-purpose electronic document management systems"* – the current state, problems, and development perspectives of medical information systems (MIS) and special-purpose electronic document management (EDMS) in the context of modern information technology application have been comprehensively analyzed.

In this chapter, initially **the specific characteristics of medical documents** were investigated. It was noted that since these documents contain critical information about individuals' health status, diagnoses, and genetic indicators, the requirements of **confidentiality**, **integrity**, **availability** (CIA triad) are set at the highest level. During the research, international standards regulating data protection (ISO/IEC 27001, NIST 800-53) and legal regulations, especially **HIPAA** (Health Insurance Portability and Accountability Act) and **GDPR** (General Data Protection Regulation) provisions

were deeply analyzed, and the compliance of existing systems with these requirements was evaluated.

During the course of the research, **threat models** targeting medical systems were classified. It was determined that traditional security mechanisms (e.g., static access control models – DAC, MAC) cannot provide flexibility for the dynamically changing medical environment (emergency situations, inter-branch exchange). At the same time, problems arising from the application of existing standard encryption algorithms (RSA, AES) were identified:

1. **Inefficiency of computational resources:** significant delays occur when processing large-volume medical images (DICOM standard MRI, CT images) with standard block encryption methods, which hinders the real-time diagnostic process.

2. **Key management problems:** Classical protocols used for secure transmission of encryption keys over open channels (e.g., Diffie-Hellman) have potential vulnerability against future quantum computations.

In the direction of solving these problems, the necessity of applying the **"Defense in depth"** strategy was substantiated. At the conclusion of the chapter, it was concluded that in order to ensure high-level protection of medical data, **deterministic chaos theory**, **fractal geometry**, **non-standard matrix algebra** and **combinatorial structures (maze theory)** the development of new, hybrid cryptographic methods and distributed database architectures based on these is a current scientific-technical issue. This approach enables both increasing cryptographic resilience and improving system performance.

In the Second Chapter, – *"Methods of Protecting Information Confidentiality in Special-Purpose Electronic Document Management Systems"* – the main scientific-technical results of the dissertation and the proposed algorithms were described.

In the section on Cryptographic Methods and Algorithms for Protecting Information Confidentiality in Special-Purpose Electronic Document Management Systems, modern cryptographic methods used to ensure information confidentiality in special-purpose, including medical, electronic document

management systems (SPEDMS), their classification, advantages, and disadvantages were extensively analyzed. During the research, it was determined that cryptographic systems applied for the protection of medical documents are divided into two main groups:

1. **Symmetric encryption systems:** here the same secret key is used for encryption and decryption (e.g., DES, AES, 3DES). Although these methods have high speed, they create the problem of secure key distribution between parties. Although the AES (Advanced Encryption Standard) standard is widely used in medical systems, it has certain limitations when processing large-volume images (real-time high-quality MRI) has certain limitations in terms of computational resource conservation.

2. **Asymmetric encryption systems:** here two mathematically related keys (public and private) are used (e.g., RSA, ElGamal, Elliptic Curves). Although these systems solve the key distribution problem, their computational complexity is many times higher than symmetric systems and they are not suitable for direct encryption of large data.

In this subsection, the differences between **stream ciphers** and **block ciphers** which have become relevant in recent times, were also investigated. Taking into account the characteristics of medical documents (large volume, high correlation, real-time transmission requirement), it was substantiated that stream encryption is more promising for this field.

Analysis of existing standards has shown that "Internet of Things" (IoT) and "Internet of Medical Things" (IoMT) devices, as well as the development of quantum computations, classical algorithms may be vulnerable against future threats. For this reason, in the dissertation, the application of cryptographic approaches based on **deterministic chaos theory** was considered necessary. The properties of chaotic systems such as **ergodicity**, **mixing** and **high sensitivity to initial conditions** properties make them suitable for cryptographic "confusion" (confusion) and "diffusion" (diffusion) principles, showing high compatibility.

Consequently, it was substantiated in this subsection that in order to ensure high security and speed in special-purpose medical

systems, in addition to traditional methods, **chaos theory**, **fractal geometry** and **combinatorial (maze) structures** there is a need for the development of new hybrid encryption algorithms based on these. This analysis served as the theoretical basis for the creation of the Brownian motion, Maze encryption, and Matrix protocols proposed in the subsequent chapters of the dissertation.

Encryption method based on the chaotic motion model of molecules for protecting text-type confidential information. In this subsection, for fast and secure encryption of text-type medical documents (epicrises, opinions, prescriptions), a new pseudo-random number generator (PRNG) and stream cipher algorithm based on the physical model of **Brownian motion of molecules** was proposed by the author. Brownian motion is the irregular, chaotic movement of microscopic particles in a liquid or gas medium. The mathematical model of this process (Wiener process) has high entropy and is extremely difficult to predict. In the proposed method, the movement of a molecule in a closed $L \times W$ space is simulated. At each step, the new position of the molecule (coordinates and velocity) is calculated using the following system of equations:

$$V_{xi} = Round\left(\left(\left((V_{xi-1} \cdot (X_i + k)) \bmod M_{max}\right), r\right) + k\right)$$
$$T_i = Round\left(\left(\left((T_{i-1} \cdot k) \bmod (T + k)\right) \bmod X_{max}\right), r\right) + k$$

Where, X_i, Y_i – are the coordinates of the molecule at the collision moment; V_{xi}, V_{yi} – are the components of the velocity vector; T_i – is the collision moment (time); * k, r – are secret parameters controlling the chaos of the system (part of the key).

Working principle of the generator: 1. Initial conditions - the initial coordinates of the molecule (X_0, Y_0), initial velocity (V_{x0}, V_{y0}) and environment parameters are entered as the key. 2. Generation: at each collision point, the fractional parts of the obtained coordinates (e.g., 48 bits after the decimal point) are taken and converted to a bit sequence. 3. Encryption: the obtained pseudo-random bit stream (Keystream) is subjected to XOR ($\oplus$) operation: $C_i = P_i \oplus K_i$ [1, 8].

The advantage of this method is its **"butterfly effect"** property: small changes in the initial parameters 10^{-15} in the initial parameters creates a completely different key stream. NIST tests have shown that the outputs of this generator fully meet the statistical randomness requirements (Figure 1).

The security of the stream cipher depends on the unpredictability of dynamically-sized secret keys. A number of standard tests were applied to evaluate the pseudo-randomness of the generated sequences.

According to the chi-square test results, the D statistic averaged in the 15-17 range, meeting the acceptance threshold of 16.9 at the α=0.05 level. From the NIST 9 monobit tests, P-values were obtained between 0.04-0.5, which satisfies the $P \geq 0.01$ condition, confirming that the sequence is random. In the key sensitivity test, completely different results were obtained during decryption by changing one bit of the key, demonstrating the high sensitivity of the cipher to the key.

Thus, the proposed pseudo-random number generator based on the Brownian motion model has been confirmed through statistical tests and key sensitivity analysis. The high sensitivity of the molecule's trajectory to initial conditions ensures the formation of a hard-to-predict key stream. The results show that this approach based on the chaotic dynamics of natural physical processes presents a promising solution for improving the security of stream encryption systems. Digital images used in medical diagnostics (X-ray, MRI, CT, USG), unlike ordinary text files, have a high correlation coefficient between pixels. That is, the color values of adjacent pixels are very close to each other, which leads to the visual structure not being completely hidden by ordinary encryption methods (e.g., block ciphers in ECB mode). To solve this problem, in the dissertation, a new hybrid encryption method based on **Depth-First Search (DFS)** algorithm-generated random maze structure and **Chaos theory** was developed. To solve this problem, in the dissertation, a new hybrid encryption method based on **Depth-First Search (DFS)** algorithm-generated random maze structure and **Chaos theory** was developed.

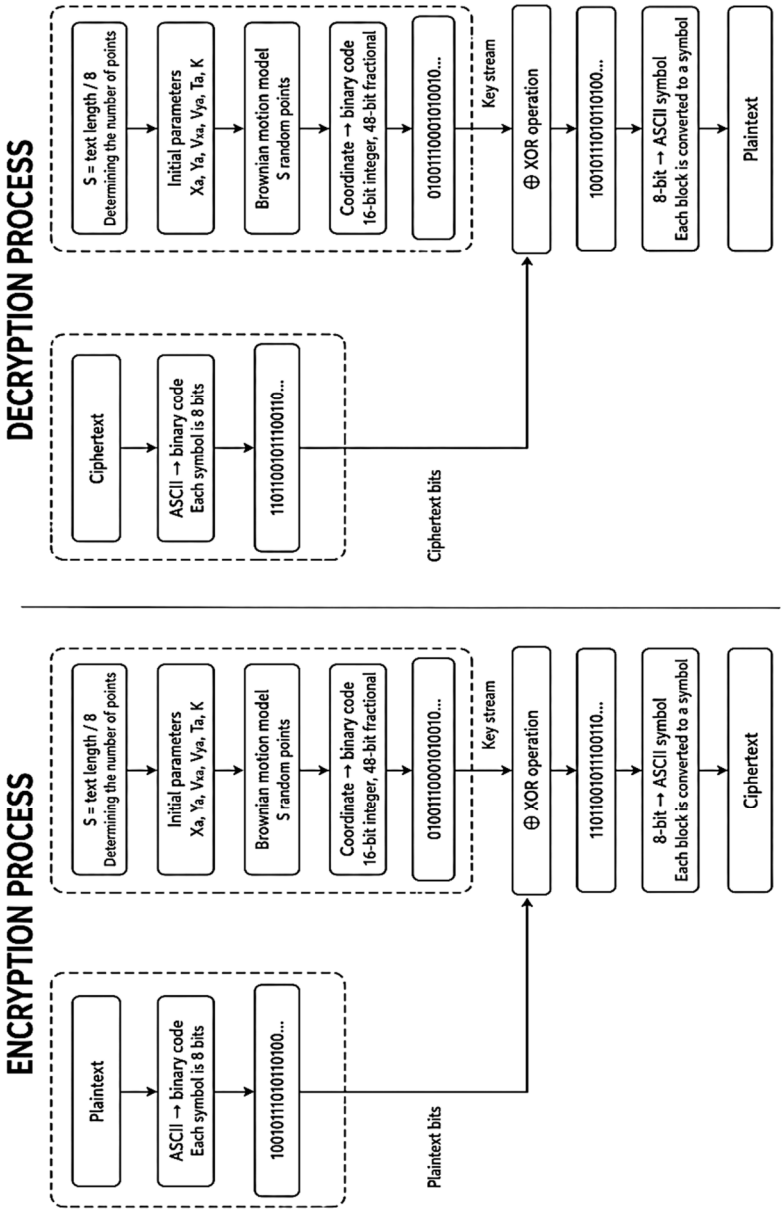


Figure 1. Schematic representation of the encryption and decryption process of algorithm

The working principle of the proposed method **"Confusion-Diffusion"** is based on the architecture and consists of the following stages:

1. Chaotic maze generation: Initially, a virtual grid of a size corresponding to the dimensions $(N \times M)$ corresponding dimensions, a virtual grid is created. The initial coordinates of the maze (x_0, y_0) and movement the selection of movement directions **Feigenbaum logistic map** as follows:

$$x_{n+1} = r \cdot x_n \cdot (1 - x_n)$$

Where $r \in [3.57,4]$ is the chaos parameter. The pseudo-random numbers obtained through this equation are used in the DFS algorithm's "neighbor cell selection" function. As a result, a maze (spanning tree) that visits all cells but has no cycles and has a unique structure is formed (Figure 2).

0	18	17	16	21	24	25	26	27	28	29	355	319	318	317	316	356	357	358	359
1	19	20	15	22	23	399	35	34	31	30	354	320	313	314	315	382	383	384	360
2	3	4	13	12	38	37	36	33	32	325	322	321	312	311	310	364	363	362	361
156	157	5	14	11	39	40	41	331	330	324	323	304	305	385	309	365	366	286	285
155	154	6	9	10	167	166	42	332	329	326	353	303	306	307	308	367	288	287	284
152	153	7	8	164	163	165	43	333	328	327	301	302	371	370	369	368	289	282	283
151	158	159	160	161	162	168	44	334	339	338	300	373	372	377	378	381	290	281	390
150	147	146	175	174	173	169	45	335	336	337	299	374	375	376	379	380	291	280	279
149	148	145	129	128	171	170	46	47	352	340	298	297	296	295	294	293	292	277	278
142	143	144	130	127	172	50	49	48	351	341	342	232	233	234	386	387	275	276	391
134	133	132	131	126	125	51	350	349	348	344	343	231	236	235	389	388	274	273	392
135	120	121	122	123	124	52	55	56	347	345	346	230	229	237	240	241	272	393	393
136	119	118	180	181	176	53	54	57	58	225	226	227	228	238	239	242	271	394	270
137	138	117	179	178	177	64	63	60	59	224	218	217	210	209	208	245	246	270	269
140	139	116	182	183	66	65	62	61	223	222	219	216	211	212	207	248	247	395	268
141	114	115	101	100	67	68	69	71	221	220	215	214	213	206	249	249	266	267	267
112	113	103	102	99	90	85	86	87	72	73	74	202	203	204	205	250	264	263	262
111	184	104	97	98	91	84	89	88	77	76	75	201	199	198	197	251	252	396	261
110	106	105	96	93	92	83	82	81	78	398	190	191	200	195	196	254	253	259	260
109	107	108	95	94	185	186	187	80	79	188	189	192	193	194	256	255	257	258	397

Figure 2. Permutation matrix

2. Construction of the permutation matrix: During the creation of the maze, the movement trajectory of the algorithm is recorded. Based on this trajectory, a unique index M **(Path Matrix)** matrix is formed. M Each element of the matrix (m_{ij}) determines the new position of the image pixels. When the pixels of the image are rearranged in this order, the visual structure of the image is completely destroyed and the spatial correlation between pixels is broken.

3. Diffusion and stack memory depth: Since permutation alone is not sufficient (because color values (histogram) do not change), the diffusion process is applied. For this purpose, the dynamically changing depth of the **Stack** memory used during the operation of the DFS algorithm is utilized. For each (i, j) cell, the number of elements in the stack is calculated and the S- stack depth matrix is created (Figure 3). The encryption process is carried out by the following XOR operation:

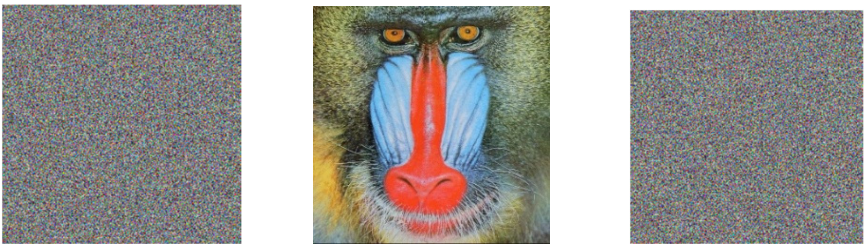
$$C_i = P_{perm(i)} \oplus S_i \oplus K_{stream}.$$

Where, C_i – is the value of the encrypted pixel; $P_{perm(i)}$ – is the permuted pixel; S_i – is the stack depth (this value changes randomly depending on the structure of the maze); K_{stream} – is the additional key stream generated by the chaotic generator.

0	17	16	15	16	19	20	21	22	23	24	207	202	201	200	199	200	201	202	203
1	18	19	14	17	18	33	30	29	26	25	206	203	196	197	198	209	210	211	204
2	3	4	13	12	33	32	31	28	27	208	205	204	195	194	193	208	207	206	205
139	140	5	14	11	34	35	36	213	212	207	206	187	188	195	192	209	210	169	168
138	137	6	9	10	145	144	37	214	211	208	209	186	189	190	191	210	171	170	167
135	136	7	8	143	142	143	38	215	210	209	184	185	214	213	212	211	172	165	166
134	137	138	139	140	141	144	39	216	221	220	183	216	215	220	221	224	173	164	163
133	130	129	150	149	148	145	40	217	218	219	182	217	218	219	222	223	174	163	162
132	131	128	119	118	147	146	41	42	229	220	181	180	179	178	177	176	175	160	161
125	126	127	120	117	148	45	44	43	228	221	222	117	118	119	178	179	158	159	162
124	123	122	121	116	115	46	229	228	227	224	223	116	121	120	181	180	157	156	163
125	110	111	112	113	114	47	50	51	226	225	226	115	114	121	124	125	126	155	164
126	109	108	119	120	115	48	49	52	53	110	111	112	113	122	123	128	127	154	165
127	128	107	118	117	116	59	58	55	54	109	104	103	96	95	94	129	130	153	166
130	129	106	119	120	61	60	57	56	109	108	105	102	97	98	93	132	131	154	151
131	104	105	92	91	62	63	64	65	66	107	106	101	100	99	92	133	148	149	150
102	103	94	93	90	81	80	81	82	67	68	69	88	89	90	91	134	147	146	145
101	104	95	88	89	82	79	84	83	72	71	70	87	86	85	84	135	136	147	144
100	97	96	87	84	83	78	77	76	73	78	77	78	87	82	83	138	137	142	143
99	98	99	86	85	86	87	88	75	74	75	76	79	80	81	140	139	140	141	144

Figure 3. XOR values table created based on stack memory

4. Iterative strengthening: The algorithm is repeated for several rounds (e.g., 5 cycles) to increase the "avalanche effect" avalanche effect. As a result, the histogram of the encrypted image falls into a completely uniform distribution, and the inter-pixel correlation coefficient approaches 0 (e.g., horizontal – 0.0021, vertical – 0.0018). These indicators prove the high resistance of the method against statistical and differential attacks (Figure 4).



a) b) c)

Figure 4. a) image encrypted with key (K1=8763572198561905), b) image decrypted with correct key (K1=8763572198561905), c) image decrypted with incorrect key (K1=8763572198561805)

Security and quality analysis. The effectiveness of the proposed algorithm was verified through large-scale experiments:

1. Histogram analysis. While the histogram (color distribution frequency) of original medical images is uneven, the histogram of images encrypted with Maze encryption falls into a completely uniform distribution. This shows that the encrypted image does not leak any statistical information about the original content and is resistant to statistical attacks (Figure 5, Figure 6).

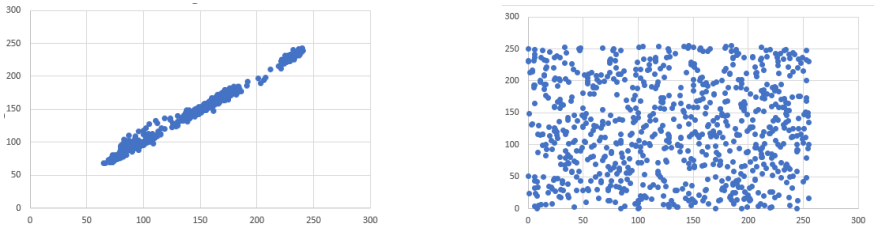
2. Correlation analysis. In the original image, the correlation coefficient between adjacent pixels is close to unity ($R \approx 0.95 - 0.98$). After applying the proposed algorithm, this indicator sharply decreased to near 0 (horizontal: 0.0021, vertical: 0.0018, diagonal: 0.0034). This confirms that the connection between pixels is completely broken and the visual structure has disappeared (Figure 7).

a) b) c)

Figure 5. Histogram analysis of original images: a) red, b) green, c) blue

a) b) c)

Figure 6. Histogram analysis of encrypted images: a) red, b) green, c) blue



a) b)

Figure 7. Correlation dependency graph of horizontal adjacent pixels of images: a) original, b) encrypted (the same result was obtained for vertical and diagonal pixels as well)

Analysis with NIST tests. The statistical indicators of the proposed algorithm were also verified through NIST tests. For this purpose, 7 tests were used and experiments were conducted on cipher-images of various sizes, with all verification results being positive (table).

Table. NIST test results

Test	Result
Frequency (monobit) test	P-value = 0.799
Frequency test within a block	P-value = 0.569
Test for the longest run of ones in a block	P-value = 0.151
Cumulative sums test	P-forward value = 0.900 P-reverse value = 0.674
Runs test	P-value = 0.998
Binary matrix rank test	P-value = 0.381
Approximate entropy test	P-value = 0.303

3. Entropy analysis. The information entropy of the encrypted image was calculated and the result **7.999** (maximum possible value is 8) was obtained. This shows that the data is completely chaotic and indistinguishable from random noise.

4. Key sensitivity. When a small change of the order of (e.g., r parameter or initial coordinate) 10^{-15} is made, is made, more than 50% of the pixels in the encrypted image change. This property ensures the high resistance of the algorithm against differential attacks.

Performance. The algorithm works several times faster than standard asymmetric algorithms (RSA) in encrypting large-volume medical images and is suitable for application in real-time systems.

Conclusion. The Maze encryption method is a strong solution both mathematically (chaos, graph theory) and statistically for ensuring the confidentiality of medical images. The algorithm renders visual data completely unrecognizable and the restoration of the original without the key is theoretically impossible [2].

As an alternative to the classical Diffie-Hellman protocol, **non-invertible (zero-determinant) matrices** a new asymmetric protocol built upon was developed. Working principle of the protocol: 1. The parties (Alice and Bob) choose a common matrix with zero determinant ($det(C) = 0$) C matrix. 2. Alice chooses the secret A matrix and the power q while Bob chooses the secret B matrix and the power p power. 3. Alice $S_A = A \cdot C^q$, while Bob chooses the $S_B = C^p \cdot B$ matrices and send them to each other. 4. The parties complete the incomplete keys they received with their own secret parameters. Final shared key: $S = A \cdot C^{q+p} \cdot B$. Even if a third party obtains S_A and $S_B$through the open channel, C since it has no inverse, cannot solve the equation and recover A and B secret matrices. This ensures the protocol's resistance against linear algebra attacks. The robustness of the protocol was analyzed for the following attack scenarios:

- **Resistance against linear algebra attacks.** The attacker obtains C, S_{A1} and S_{B1} matrices from the open channel. The goal is to find the secret A or B matrices. However, $S_{A1} = A \cdot C^q$ in the

equation C matrix **non-invertible** so this system of equations becomes an **underdetermined** system. That is, the number of equations is less than the number of unknowns and an infinite number of A matrices can satisfy this equation. The attacker cannot uniquely determine the true A matrix. Additionally, the matrices introduced in the second round M and N matrices further increase the complexity of the equation system and exclude trivial solutions.

- **Key space and "Brute force" analysis.** The resources required to break it are measured by the size of the key space (number of possible matrices). Combinatorial calculations conducted in the dissertation show that $[0,1000]$ for just 4×4 matrices with elements in the given range, the number of non-invertible matrices is approximately 10^{60} in order. Bu, 200-bit encryption key and makes the "exhaustive search" (brute-force) method impossible even for modern supercomputers. As the matrix size (n) increases, the security level increases exponentially.
- **Man-in-the-middle attack.** When the proposed protocol is used together with digital signatures or certificates, M and N it demonstrates resistance against active eavesdropping attacks due to the matrices being known only to the parties.

The operating speed of the protocol was compared with classical algorithms. The **modular exponentiation** operation used in the Diffie-Hellman protocol requires significant computational resources for large numbers (e.g., 2048 bits). In the proposed method, the main operation is **matrix multiplication**. Matrix multiplication operations can be highly parallelized on modern processors (CPU/GPU). Experiments conducted (in an Intel Core i5, 8GB RAM environment) showed that the key generation and exchange process 100×100 matrices with elements in the given range, the number of 0.23 seconds. At the same security level, the matrix-based protocol showed results approximately 3-5 times faster than traditional RSA and DH algorithms.

These features prove that the proposed protocol has high potential for application in environments with limited computational

resources (IoT medical sensors, mobile terminals) and real-time systems [3, 7, 9-10].

During the analysis of large-volume data ("Big Data") collected in healthcare systems for scientific research and statistical purposes, the privacy of personal data (patient identity) must be ensured. However, traditional fully random shuffling methods (e.g., Fisher-Yates shuffle) completely destroy the structure of data, which reduces their analytical utility. In the dissertation, to solve this problem, **Maze Shuffling** based on the principle of local randomness, was developed.

The main purpose of this method is to shuffle data in such a way that re-identification of individual records is impossible, while the general statistical properties of the database (local clusters, distribution patterns) are preserved.

Working principle of the algorithm: 1. **Spatial representation of data:** A 1D (one-dimensional) data array (e.g., patient list) is transformed into a 2D (two-dimensional) virtual matrix or maze ($N \times N$) transformated. 2. **Maze trajectory:** A random path is created using the DFS algorithm. Data displacement is carried out along this path. 3. **Lazy walk mechanism:** The main novelty of the algorithm is the application of the **"Lazy Walk"** model. For each iteration (t) the probability of the element swapping places with a neighboring cell (p_{swap}) the probability of staying in place (p_{lazy}) is also taken into account. If $Rand_{val} < p_{lazy} \rightarrow$ the element stays in place, if $Rand_{val} \geq p_{lazy} \rightarrow$ the element swaps places with its neighbor in the maze. This approach ensures that the system is **aperiodic** and, according to Markov chain theory, guarantees that as the number of iterations increases, the system approaches a stable mixing state.

Mathematical justification. It was proved in the dissertation that Maze Shuffling algorithm's transition matrix (K) **primitive** (irreducible and aperiodic). According to the Perron-Frobenius theorem, this means that the number of iterations (t) increases, the distribution of data approaches the uniform distribution at an exponential rate. The privacy risk (R_t) is evaluated as follows:

$$R_t = \| K^t - J \| \leq (\sigma_2(K))^t.$$

Where $\sigma_2(K)$ is the second largest singular value of the transition matrix (< 1). This formula shows that as iterations increase, the connection of personal data with its initial position (re-identification risk) approaches zero exponentially.

Conclusion. The Maze Shuffling algorithm achieves an "Fisher-Yates" (full chaos) and "Block-Shuffling" weak confusion methods, optimal balance. It sufficiently shuffles data to support **k-anonymity** and **l-diversity** principles, while preserving the locality property of data and remaining useful for medical analyses (e.g., geographical spread of diseases or clustering by age groups). This method is especially ideal for ensuring security when processing and sharing large-volume medical data in cloud environments with researchers [12].

In the Third Chapter, *"Organizational and technical aspects of ensuring information security in special-purpose electronic document management systems"* In this chapter of the dissertation, the system architecture, distributed database structure, and access control mechanisms necessary for the application of the proposed cryptographic methods and protocols in real medical information systems (MIS) were developed.

During the research, it was determined that modern medical institutions (hospitals, polyclinics, laboratories) usually have a complex structure consisting of a central office and geographically dispersed branches. In such systems, centralized data storage creates the "Single Point of Failure" (Single Point of Failure - SPOF) risk and paralyzes branch operations during network disruptions. To solve this problem, in the dissertation, a new hybrid encryption method based on **hybrid distributed architecture** was proposed. In this model, each branch has its own local database but is simultaneously synchronized with the central server (cloud or data center).

The main components of the security model are as follows:

Electronic signature infrastructure. An improved electronic signature scheme is applied to ensure the integrity and non-

repudiation of medical documents (prescriptions, epicrises). In this scheme, the hash function of the document (e.g., SHA-256) is calculated and encrypted with the author's secret key.

Dynamic roles and permissions. A dynamic RBAC model was developed to manage the authorities of medical personnel (doctor, chief physician, laboratory technician, registrar) in the system. In this model, in addition to static roles, depending on the situation (e.g., on-call doctor mode, emergency) temporary permissions can be granted or revoked [5].

To ensure the uninterrupted operation of medical institutions, in the dissertation, a distributed database architecture based on the **leaderless replication** model was proposed. Unlike the traditional Leader-Follower architecture, in this model each branch (node) can independently execute both read and write operations.

Working principle of the proposed architecture: 1. **Local write:** The doctor writes the document to the branch's local database (works even in offline mode). 2. **Asynchronous replication:** When the network connection is restored or periodically, local changes are transmitted to other branches and the central database. 3. **Fragmentation:** Data is distributed based on both horizontal (patient groups by branch) and vertical (by data type - text, image) fragmentation principles, which increases query speed.

The biggest problem of leaderless systems is the **conflicts**. that arise when changes are made to the same document at different nodes simultaneously. For example, the doctor at branch A changes the patient's diagnosis while the laboratory technician at branch B adds the analysis results of the same patient at the same time. To solve this problem, in the dissertation, a new synchronization algorithm based on **"Transaction table"** and **timestamp** was developed:

- **Version vectors.** A version vector is maintained for each document. The vector is incremented upon changes.
- **Transaction table.** A special table is created at each node that stores the status (executed/pending) of the executed transactions.
- **Conflict resolution:** During data merging between two nodes, hash-structured sets are created by object IDs:

$$Tr_{conflicts} = HashSet_{by_{ObjectID}}\left(\sum Tr_i^A + \sum Tr_j^B\right).$$

The system checks timestamps for conflicting objects. “Last writer wins” principle or semantic merging (if changes are in different fields) is applied so that the most current version is propagated to all databases, while the other is rolled back.

This approach ensures that the system is **the CAP theorem** high availability and partition tolerance in accordance with [6].

In the Fourth Chapter, *“Software solutions and application of methods and algorithms for protecting information confidentiality in special-purpose institutions”* the software-level implementation of the theoretical methods and algorithms proposed in the dissertation, their operating principles, statistical analysis of the obtained results, and comparison with existing standards were presented. The developed cryptographic methods were implemented in the **C#** programming language, in the *Microsoft Visual Studio* environment. The created software complex consists of the following functional modules:

Text encryption module. Encryption and decryption of text-type documents (TXT, DOCX) at the bit level using XOR operation by generating pseudo-random numbers based on the Brownian motion model of molecules, according to initial parameters (coordinates, velocity) entered by the user.

Image encryption module. Encryption of image files (BMP, JPG) through the DFS maze algorithm using maze generation with initial values via the Feigenbaum and DFS methods and permutation and value change methods through two matrices resulting from this generation. The encryption process, visual representation of the maze, and the result of the encrypted image (“noise” pattern) are displayed in the user interface.

Maze Shuffling module. Application of the shuffling method using Python software in CSV-type documents for preserving the anonymity and utility of tabular data with a matrix generated via DFS.

Key exchange module: Simulation of the matrix algebra-based protocol. Here, the process of matrix generation between

parties, computation of products, and obtaining the shared key was automated in the Python language.

Analysis of obtained results and development of recommendations

The advantages and favorable application areas of the methods and algorithms proposed in the dissertation were analyzed as follows:

- **Uniqueness of the randomness source.** The method is inspired by a physical process (chaotic motion of molecules, even if simulated) for the pseudo-random number generator (PRNG). This is a different approach from traditional chaotic PRNGs based purely on mathematical functions and proposes a new type of randomness source.
- **Use of chaotic properties.** The chaotic nature inherent in Brownian motion, namely high sensitivity to initial conditions and unpredictability, is the basis for increasing resistance to cryptanalysis. A small change in the key (initial coordinates, velocities, k coefficient) creates a completely different pseudo-random number sequence (encryption stream).

Since the proposed algorithm is generally a stream cipher method, it can be widely used in areas where stream encryption is applied. These properties make it suitable for the following areas:

1. **Real-time communication.** Systems where data comes as a continuous stream and low latency is essential.
2. **Resource-constrained environments.** Devices with low computing power, memory, or energy consumption.
3. **High-speed data transmission.**
4. **Cases where error propagation needs to be minimized.**
5. **Ease of software implementation.**

Based on application areas, it can be generalized:

- **Military and defense structures:** ensuring real-time, secure communication is critical.
- **Intelligence and counter-intelligence agencies:** establishing confidential communication channels, real-time transmission of operational data.

- **Law enforcement agencies and emergency services:** ensuring secure and uninterrupted voice data communication during operations.
- **Diplomatic corps and ministries of foreign affairs:** secure transmission of confidential data between the domestic center and embassies and consulates abroad.

The proposed maze-based image encryption method for encrypting image-format documents presents the following potential contributions and advantages to the field of image cryptography and general security:

- **Consideration of image characteristics:** the method is specifically designed for encrypting images.
- **Dual cryptographic mechanism:** the method consists of two main steps: pixel shuffling and pixel value modification.
- **Combination of chaotic and random processes.** The Feigenbaum chaotic function (for PRNG) and random selections of the DFS algorithm are used as keys for maze generation.
- **Confirmation through security analysis.** The analyses conducted in the work show that the method is resistant to statistical attacks and differential analysis.
- **Strengthening through repeated operations.** Repeating the shuffling and replacement steps at least 5 times serves to strengthen resistance to cryptanalysis by increasing the effect of diffusion and confusion principles.

Advantages of anonymization through shuffling:

- Preservation of statistical properties (at column level).
- Conceptual simplicity.
- Ease of implementation.
- Breaking the connection: shuffling breaks the direct connection between values in different columns within the same row (record).
- Low data loss (for column analysis).

Potential advantages of the matrix-based key exchange protocol:

- **Different mathematical foundation:** the protocol is based on a mathematical problem different from factorization and discrete logarithm.
- **Potential performance advantages (in specific cases):** matrix operations can be very efficiently executed on modern processors or specialized hardware.

In general, there are many important points for the security of information technology infrastructure in every electronic document management system that should be applied and taken into account in every modern IT corporate environment. These important points are listed below:

- it is essential to always update applications used on computers to new versions.
- not opening suspicious items received via email.
- using antivirus software.
- every connection should be established through a secure channel using VPN protocols.
- it is essential to fulfill the required conditions for passwords to ensure the security of each user's password entering the system.
- using two-factor authentication.
- every user using the system should be informed about not visiting suspicious websites or links.
- Programmers should implement validations and restrictions for every parameter in queries entering the system to prevent SQL injection. For system additions on web applications, another threat to consider is CSRF attacks.
- passwords of every user using the system should not be stored in databases in plain text.
- buffer overflow. For its solution, validation of input data or providing an error message during overflow is essential.
- denial of service. To prevent such cases, external queries should be distinguished by the server to prevent server crashes.
- broken access control: To solve this deficiency, permissions should be correctly created for each user based on the principle

of least privilege, and access to resources should be checked with stricter conditions before granting permissions.

In conclusion, the tests and comparative analyses conducted in the dissertation confirmed that the application of the proposed methods in medical information systems is efficient in terms of both security and performance.

CONCLUSIONS

1. Strengthening cryptographic systems through the application of chaotic processes. In the research, the security of cryptographic methods was enhanced by utilizing the random-like and hard-to-predict nature of chaotic processes. This approach, by creating dynamic and complex structures especially in key generation and encryption processes, has made it possible to eliminate the weaknesses of existing algorithms and has formed a more reliable and resilient foundation against external attacks for cryptographic applications.

2. New stream encryption algorithm for text-format documents. As a scientific novelty, a new stream encryption algorithm operating with a random number generator based on the chaotic Brownian motion model of molecules was created for encrypting text-type medical data. This algorithm applies unique encryption for each character by combining the generated random bit sequence with the bits of the text. The proposed method provides high-speed encryption while also demonstrating high resistance to statistical analyses.

3. Maze-based encryption for image-format data. A new encryption method using mazes generated by the Depth-First Search (DFS) algorithm was developed for effectively protecting the content of medical images. Generation This method shuffles the pixels of the image along the paths of the maze, making it extremely complex to restore their original positions, renders the visual information of the images completely unrecognizable, and makes the process of restoring the original image without the key theoretically impossible.

4. **Key exchange protocol based on matrix algebra.** A new protocol based on the one-way functions of matrix algebra (non-invertible matrices) was developed to ensure secure key exchange between parties during data transmission in networks. This protocol enables the parties to obtain a common secret key through open channels and ensures privacy through matrix operations that are difficult to reverse-compute.

5. **New algorithm for data anonymization.** An anonymization algorithm based on the shuffling matrix obtained from maze generation was created to ensure the protection of individual privacy when using data for analytical purposes. This algorithm hides the identity of sensitive data in databases by shuffling them in an irreversible manner and enables analysts to perform analyses on data without violating individual privacy.

6. **Synchronization and replication architecture for distributed databases.** A new synchronization and replication architecture based on transaction tables was developed to ensure the integrity and consistency of data in distributed medical systems. This model records all operations performed on different copies of the database in a special table, ensuring the orderly transfer of changes to other copies, and prevents data loss and conflicts.

7. **Creating protected channels for secure document transmission.** A method for creating reliable communication channels for securely transmitting documents to remotely located databases was proposed. This method protects transmitted data with strong encryption algorithms and ensures the isolation of data from external interference using tunneling technologies.

8. **Dynamic authentication and authorization models for resource access.** An authentication and authorization model based on a dynamic roles and permissions system was developed to ensure access control to resources in information systems. This model, in addition to verifying users' identities, determines which data and functions they can access according to their roles in the system. The proposed dynamic approach increases security and management flexibility by enabling the management of roles and permissions in real-time.

The main results of the dissertation have been published in the following scientific works published:

1. Gasimov, V.A., Mammadov, J.I., Mammadzada, N.F. Stream encryption method based on the chaotic brownıan motion model of molecules // 4th International Conference on Innovative Data Communication Technologies and Application (ICIDCA 2022). 3-4 November, http://icidca.com/2022/. – Coimbatore, India. Procedia Computer Science, vol.215, – 2022, – pp. 577-588.
2. Gasimov, V.A. Maze based image encryption method constructed by random number generation / V.A.Gasimov, N.F.Mammadzada, J.I.Mammadov [et al.] // Eurasian Journal of Mathematical and Computer Applications. ISSN 2306-6172. Vol. 12, Issue 3, – 2024, – pp.35-50. (Web of Sciences).
3. Gasimov, V.A., Mammadzada, N.F., Mammadov, J.I. New Key Exchange Protocol Based on Matrix Algebras // 5th International Conference on Problems of Cybernetics and Informatics (PCI), Azerbaijan, – Baku, – 2023, – pp. 1-3, doi: 10.1109/PCI60110.2023.10326004.
4. Gasimov, V.A., Mammadov, J.I., Mammadzada, N.F. Method of hidden transmission of information based on fractals and its software // Tomsk State University, Journal of Control and Computer Science, – Russian Federation. – 2023(65), – pp.95-104. DOI: doi: 10.17223/19988605/65/10
5. Gasimov, V.A., Mammadzada, N.F., Mammadov, J.I, Aliyeva, K.J. The problem of information protection in electronic document management systems of medical organizations // Problems of Informatization and Management. – 2025, – pp.37-45. 10.18372/2073-4751.81.20127.
6. Gasimov, V.A., Mammadzada, N.F. Architecture and model of a special-purpose electronic document management system with a distributed structure // Mathematics and Computer Science, – 2024, vol. 8, no.2, – pp. 120-129.
7. Gasimov, V.A., Mammadzada, N.F., Mammadov, J.I. Matris əsaslı yeni açar mübadiləsi protokolu // II International

Conference on Informatıon Security: Problems and Prospects, – Baku, – 2022.
8. Gasimov, V.A., Mammadzada, N.F., Mammadov, J.I. Molekulların xaotik hərəkətinə əsaslan yeni simmetrik şifrələmə alqoritmi // II International Conference on Informatıon Securıty: Problems and Prospects, – Baku, – 2022.
9. Gasimov, V.A., Mammadzada, N.F., Mammadov, J.I. // Investıgatıon Of The Securıty Of A Key Exchange Protocol Based on Matrıx Algebra // Mathematics and Computer Science, – 2024, vol. 8, no. 2, – pp.101-104.
10. Mammadzada, N.F. Matrıx Based Key Exchange Protocol // Proceedıngs of the 8th International Conference on Control And Optimization With Industrial Applıcations. – Baku, – 2022.
11. Gasimov Vagif, Mammadzada Nargiz. Data Shuffling Algorithm For Preserving Privacy and Utility Of Analysis, ITTA 2026 (3rd International Conference on Information Technologies and Their Applications), Azerbaijan
12. Gasimov A. Vagif, Mammadzada F. Nargiz, Mustafayeva A. Esmira, Aliyeva J. Kamala, Probabilistic stochastic shuffling for privacy-preserving data analytics with maintained analytical utility and robust statistical performance, Advanced Mathematical Models & Applications (in press).

The defense will be held on 30 June 2026 at 1400 at the meeting of the Dissertation council FD 2.48 of Supreme Attestation Commission under the President of the Republic of Azerbaijan operating at Azerbaijan State Oil and Industry University.

Address: AZ1010, Baku, Azadlig avenue 20, Azerbaijan State Oil and Industry University.
E-mail: info@asoiu.edu.az

Dissertation is accessible at the library of the Azerbaijan State Oil and Industry University.

Electronic version of the abstract is available on the official website of the Azerbaijan State Oil and Industry University.

Abstract was sent to the required addresses on 27 May 2026.

Signed for printing: 26.05.2026

Paper format: A5

Volume: 38595

Number of hard copies: 20