Wireless Networks
https://doi.org/10.1007/s11276-023-03436-z

ORIGINAL PAPER

Advancing database security: a comprehensive systematic mapping study of potential challenges

Asif Iqbal[1] · Siffat Ullah Khan[1] · Mahmood Niazi[2,3] · Mamoona Humayun[4] · Najm Us Sama[5] · Arif Ali Khan[6] · Aakash Ahmad[7]

Abstract

The value of data to a company means that it must be protected. When it comes to safeguarding their local and worldwide databases, businesses face a number of challenges. To systematically review the literature to highlight the difficulties in establishing, implementing, and maintaining secure databases. In order to better understand database system problems, we did a systematic mapping study (SMS). We've analyzed 100 research publications from different digital libraries and found 20 issues after adopting inclusion and exclusion criteria. This SMS study aimed to identify the most up-to-date research in database security and the different challenges faced by users/clients using various databases from a software engineering perspective. In total, 20 challenges were identified related to database security. Our results show that "weak authorization system", "weak access control", "privacy issues/data leakage", "lack of NOP security", and "database attacks" as the most frequently cited critical challenges. Further analyses were performed to show different challenges with respect to different phases of the software development lifecycle, venue of publications, types of database attacks, and active research institutes/universities researching database security. The organizations should implement adequate mitigation strategies to address the identified database challenges. This research will also provide a direction for new research in this area.

Keywords Database security · Systematic mapping study · Secure databases · Modeling and maintenance of protected databases · Issues in the development

1 Introduction

Companies' databases (DBs) are repositories of their most significant and high-value data. As DB utilization has surged, so has the frequency of attacks on these databases. A DB attack is characterized as an event that jeopardizes a resource by altering or destroying vital data [1, 2]. The common goal of DB attacks is to access critical information. Illicitly acquiring sensitive data such as credit card details, banking data, and personal identifiers is another prevalent motive behind DB hacks. In our interconnected global society, several technologies provide avenues for

✉ Arif Ali Khan
arif.khan@oulu.fi

[1] Department of Computer Science and IT, Software-Engineering-Research-Group (SERG-UOM), University of Malakand, Chakdara, Pakistan

[2] Information and Computer Science Department, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia

[3] Interdisciplinary Research Center for Intelligent Secure Systems, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia

[4] Department of Information Systems, College of Computer and Information Sciences, Jouf University, Sakaka 72311, Saudi Arabia

[5] Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak, Kota Samarahan, Sarawak, Malaysia

[6] M3S Empirical Software Engineering Research Unit, University of Oulu, Oulu, Finland

[7] Lancaster University Leipzig, Leipzig, Germany

Published online: 17 July 2023

DB attacks to exploit vulnerabilities in DB architecture, as per common understanding [1, 3, 4].

Many enterprises confront challenges like data piracy, data replication, and denial of service attacks. To infiltrate a company's DBs, cybercriminals scout for system vulnerabilities and exploit them using specialized tools [5, 6].

The aspect of security should be prioritized during the development of information systems, particularly DBs. In terms of software development, security concerns must be addressed at every stage of the development cycle [7]. As illustrated in Fig. 1, security breaches, including the loss of critical data, have become commonplace in recent years. Given the importance of data security to numerous businesses, a range of measures and methodologies are required to safeguard the DB [8–10]. A secure DB is designed to react appropriately in the event of a potential DB attack [11].

In the current world, the impact of cyber-attacks on the commercial landscape must be addressed. To succeed in the globalized environment, businesses must ensure the protection of their vital data. DBs can be safeguarded from unauthorized access [12–14]. When a DB is outsourced to the cloud, cloud platforms introduce security challenges such as unreliable service providers, malicious cloud employees, data protection, consistency, and scalability. With cloud DBs becoming increasingly susceptible to both external and internal threats, traditional and conventional security measures are insufficient for their protection [15, 16].

While extensive work has been done in this field, much of it focuses on a few specific DB platforms or problems, typically explored through standard literature reviews. We aim to provide a more holistic view by conducting a systematic mapping study (SMS) to identify security concerns in DB architecture, development, and maintenance from a software engineering perspective. This SMS will help us identify the ongoing research challenges and priorities.

The following research questions (RQs) will guide our SMS to achieve our study objectives:

RQ1 What is the current state of the art in the development and implementation of secure DBs?

RQ2 What are the security issues in building, implementing, and maintaining secure DBs, as reported in the literature?

1.1 Paper contribution

The contributions of the intended work are as follows:

- The proposed research undertakes a systematic mapping study (SMS) to identify and emphasize the challenges associated with developing and maintaining secure databases.
- In addition to showing the difficulties experienced by users using various databases from a software engineering standpoint, our SMS survey sheds light on some of the most current database security studies.
- It also highlights the importance of maintaining careful attention to database security and suggests a direction for future research in this field.

1.2 Motivation for the paper

Several research in the literature seeks to give a solution for database security. However, before moving forward with new solutions, it is necessary to synthesize current knowledge to offer security practitioners the most up-to-date information. We must identify the cutting-edge in constructing, implementing, and maintaining dependable databases, as well as security challenges, so BD's design, development, and maintenance may be secure. The

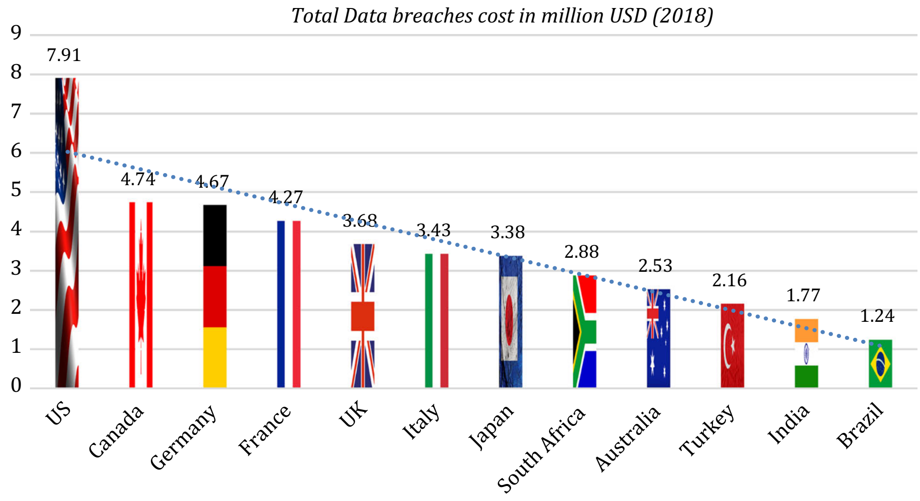


Fig. 1 Total data breaches cost in different countries [5]

motivation behind this research is to provide in-depth solutions to these problems.

1.3 Paper organization

The remainder of the article is arranged in the following manner.

In Sect. 2, we discussed the background of DB security, and Sect. 3, illustrated the research methodology in detail. The results of our conducted SMS are given in Sect. 4. In Sect. 5, the Implication of our findings is discussed. Finally, the conclusion and future work are discussed in Sect. 6. Other supportive information is provided in the rest of the sections at the bottom of this paper.

2 Background

There are a number of studies that look at database security from different angles. In their study [17], Mai et al. suggest using cloud-based security measures to safeguard power system databases. Using an RSA encryption method, public and private keys are generated for database encryption; a huge prime integer is chosen randomly from the cloud platform's Simple Storage Service and used as the client key. When the database receives a verification key, it compares it to the public key and private key established by the RSA encryption method. If the database determines that the access is legitimate, it provides feedback on the access. According to the findings of the tests, the database can be protected against threats as the threat situation value is always less than 0.50 once the design technique has been implemented.

A data encryption algorithm was developed by Ibrahim et al., which provides an encryption-based solution for DB security. In this system, information is encrypted using standard ASCII characters. They encrypted all of the data in the database and used three keys to access the primary formula. Numbers and text both work for the data. The suggested formula may restore the data's original format by combining another coordinator with the aforementioned three keys. In order to achieve a comparable data size to when the data is encrypted at a decent pace, the algorithm prioritizes data size and recording speed [18].

The article offers a lightweight cryptosystem based on the Rivest Cypher 4 (RC4) algorithm [19] as a solution to the widespread problem of insecure database transfer between sender and recipient. This cryptosystem safeguards sensitive information by encrypting it before sending it through a network and then decrypting it upon its safe return. Database tables have an encapsulating system that ciphers symbolize hens.

The continual improvements in digitizing have enhanced the prominence of online services. Enterprises must store essential data in corporate DB systems, including bank records, activities, the history of patient paperwork, personal data, agreements, etc. The institutions also must maintain the data's authenticity, privacy, and availability. Any intrusion in security procedures or data may cause severe economic loss and damage the company's reputation [20]. The remarkable development in the deployment of DB's is the required architecture to cope with information that can be attributed to the rising big data. Every 1.2 years, according to research, the entire quantity of institutional information doubles [21].

Most of the latest studies provide encryption-based solutions for DB security. However, before proceeding towards these solutions, there is a need to find out the flaws that lead to security breaches.

One or more of the following sources can lead to a security flaw:

Interior Internal origins of attack originate from inside the corporation. Human resources—organization supervisors, admins, workers, and interns—all fall within this category of insiders. Almost all insiders are recognized in a particular way, and just a few IT professionals have significant access levels.

Exterior Exterior attacks originate from entities outside the organization instance, cybercriminals, illegal parties of established ways, and government agencies. Usually, no confidence and trust, or benefit is offered for external sources.

Collaborator Any third party involved in a business connection with the organization, firm, or group is considered a partner in many companies. This significant collection of partners, distributors, vendors, contractual labor, and customers is known as the entire enterprise. There must be some level of confidence and privilege of accessibility or record among colleagues in the entire enterprise; therefore, this is often advised.

2.1 Secure databases

With incredibly high secure data and an expanded online presence, the worries concerning DB security are high at all-time. As more systems are connected and brought online to improve access, the sensitivity towards attacks is also increased, estimated to be about $1.3 million in massive financial losses; these mischievous attacks are also liable for public reputation and client relations with the association [21, 22]. All users can boundlessly get information from the DB server in an un-secure DB system. All hosts are allowed to associate with the server from any IP address and link with the DB server, making everyone's information accessible in the storage engine [23, 24].

Hence, the DB system is retained with numerous security mechanisms which contain anticipation of unauthorized access to data from an insider or outsider of an organization. Proper encryption techniques should be applied to secure the DB's [25]. The most comprehensive secure DB model is the multilevel model, which allows the arrangement of information according to its privacy and deals with mandatory access control MAC [7]. DB services are intended to ensure that client DB's are secure by implementing backup and recovery techniques [26].

The DB can be protected from the third party, which is not authorized by the procedure called cryptography and utilizing other related techniques. The primary motivation behind DB security is ensuring data privacy from unauthorized outsiders. The essential techniques in DB security are authentication, confidentiality, and integrity, which are utilized to secure the DB's [27]. DB construction, in particular, must consider security as the main goal while developing a data system. In this respect, security should be addressed at all stages of the software development process [7, 28–30].

2.2 Related work

Various articles examine the importance of security controls from the perspective of software engineering [31]. For instance, MÁRQUEZ et al. [32] conducted a systematic survey concentrating on the telemedicine platform's safety from the software engineering viewpoint. The key focus of this article is investigating how Software development assists in designing a reliable telehealth platform. However, the proposed work is just restricted to, particularly telehealth systems.

Al-Sayid et al. [1] notably studied the challenges of data stores and proposed DB security issues. To prevent unauthorized access to or alteration of the DB's critical material, they observed a wide variety of DB security issues. Another research by Zeb focuses on identifying potential attacks on the DB system using a standard research study. Mousa et al. [33] discover the various risks to DB safety in their analysis through the unstructured research study. Moghadam et al. [15] did an investigation on cloud servers to figure out all conceivable threats.

Nevertheless, this analysis is solely restricted to the cloud DB environment. The researchers Segundo Toapanta et al. [5] uncovered real-world examples of cybercrime. Apart from that, their research is restricted to cyberattacks.

The authors in [21] have suggested an innovative technique for spotting distinct threats to DB systems by assessing the risk for incoming new activities. Their research discovered various harmful attacks that could harm the DB system. The emphasis of their research is only confined to security assessment involving DB's. Experts in [32] present a comprehensive mapping analysis, and their observations are only limited to the Telehealth system's privacy from the software engineering point of view. They did not define the security problems in creating, implementing, and managing safe DB's. Furthermore, with the rapid development of ICTs, it is essential to be up to date on the most recent developments in this field.

The primary goal of this research is to gain a greater understanding of this topic by conducting a Systematic Mapping Survey to identify the problems in building, managing, and sustaining reliable DB's.

3 Research methodology

The goal of this study was accomplished by evaluating the current state of DB privacy and suggesting areas that needed further research work. With the SMS, researchers may better connect the data from literary research to a series of questions [34, 35]. SMS is a descriptive investigation that involves picking and putting combine all published research articles associated with a particular challenge and gives a broad summery of existing materials relating to the particular questions. In the near future, software engineers will benefit significantly from SMS because it provides a comprehensive overview of the research in the field. Figure 2 outlines the process that was followed to conduct the mapping study.

3.1 Research questions

Our primary objective is to find the obstacles in planning, creating, and managing data protection. To achieve this objective, relevant study questions have been devised.

*RQ 1*What is the current state of the art in the development and implementation of secure DBs?

To address RQ1, we have studied the material depending on the sub-questions mentioned above:

*RQ 1.1*n terms of reliable data modeling, development, and maintenance, which stage has received the most attention in the research?

*RQ 1.2*What are the primary sites for robust DB design?

*RQ 1.3*What are the ongoing research organizations working in robust data modeling?

*RQ 1.4*What kinds of DB attacks have been described in the research?

*RQ 1.5*According to the research, what are the various categories of DB's?

RQ 1.6 What kinds of DBMS platforms are often employed, as stated in the literature.

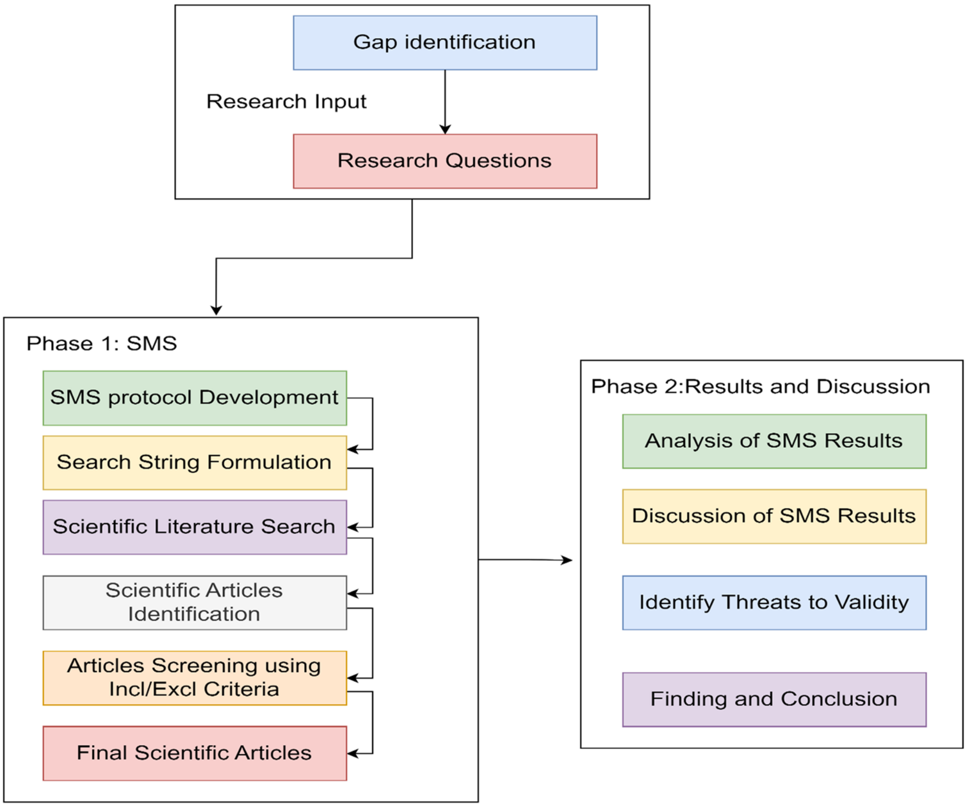

Gap identification
Research Input
Research Questions
Phase 1: SMS
SMS protocol Development
Search String Formulation
Scientific Literature Search
Scientific Articles Identification
Articles Screening using Incl/Excl Criteria
Final Scientific Articles
Phase 2:Results and Discussion
Analysis of SMS Results
Discussion of SMS Results
Identify Threats to Validity
Finding and Conclusion

Fig. 2 SMS process

*RQ 2*What are the security issues in building, implementing, and maintaining secure DBs, as reported in the literature?

3.2 Search strategy

The scholars in [36–38] employed the PICO (Population, Intervention, Comparison, and Outcomes) framework to develop a list of terms and then drew search terms from research questions.

Population DB's and software development in general.

Intervention Security Strategies.

Comparison No assessments proceed for the ongoing investigation.

Outcomes Reliable DB's.

3.3 Search strings

After several tries, the following two search terms were selected to link the PICO aspects by utilizing Boolean connector (AND):

> ((“Database security” OR “Secure Databases” OR “Database protection” OR “Guarding Database” OR “Database intrusion” OR “Database prevention”) AND (“Security Mechanisms” OR “Security Models” OR “Security methods” OR “Security policies” OR “Security techniques” OR “Security Guidelines”)).

For Science Direct online repository, we compressed the above search term due to space limits. As a result, the accompanying keywords were entered into the ScienceDirect database:

> ((“Database security” OR “Secure Databases” OR “Database protection” OR “Guarding Database” OR “Database prevention”) AND (“Security Mechanisms” OR “Security methods” OR “Security techniques” OR “Security guidelines”)).

3.4 Literature resources

We choose below digital repositories (A to F) to do our SMS and execute the search stings for acquiring publications.

- ACM–A
- IEEE xplore–B
- Springer link–C
- AIS electronic library (AiSel)–D
- Science direct–E
- Wiley online library–F

3.5 Research evaluation criteria

Titles, abstracts, entire readings, and quality assessments were all factors in our selection of research publications. The primary goal of the selection process is to compile an appropriate collection of papers by imposing inclusion and exclusion standards on submissions. We have set the accompanying inclusion and exclusion criteria to perform our SMS effectively. The same inclusion and exclusion criteria have been used in other studies [39–41]

3.5.1 Inclusion criteria

Only articles that meet one or more of the below criteria were considered for inclusion in our collection.

I1 Research involving the design and implementation of database security measures.

I2 Research that explains how to protect DB's.

I3 Research the difficulties and dangers of creating, implementing, and maintaining safe DB's.

I4 Research on the planning, development, and management of reliable DB's included in this category.

3.5.2 Exclusion criteria

The preceding exclusion criteria were considered to find relevant articles.

E1 Publications that are not published in the English language.

E2 No consideration will be given to materials that haven't been published in any journal, magazine, or conference proceedings, such as unpublished books and grey material.

E3 Books as well as non-peer-reviewed articles, including briefs, proposals, keynotes, evaluations, tutorials, and forum discussions.

E4 Articles that aren't published in their whole digital.

E5 Publications that don't meet the inclusion requirements.

E6 Research is only provided as abstracts or PowerPoint slides.

We used the snowballing approach [42–44] in addition to the previous inclusion/exclusion criteria for our concluding decision. The snowball method was used to choose seven articles from various research repositories. Appendix 1 contains the papers selected using the snowballing approach, from 94 to 100. In the latest research, scholars have employed the same method [45, 46].

3.6 Quality evaluation

All articles chosen in the selection have been evaluated for quality. Criteria for quality evaluation include:

To evaluate the papers, we used a three-point Likert scale (yes, partially, no) for every element of the quality evaluation criteria. We awarded each element of quality assessment criteria a score of 2 (yes), 1 (partially), or 0 (no) to achieve notable findings. Including an article in the SMS is permitted if it gained an average standard score of > or = 0.5. Many other scholars [45, 47–49] have employed a similar approach. A list of all of the questions from Table 1 is included in the quality ranking.

3.7 Article selection

Employing Afzal et al. tollgate's technique, we adjusted the key publication selection in our SMS analysis upon executing the search terms (Sect. 3.3) and online DB's (Sect. 3.4) [50]. The five stages of this method are as follows: (Table 2).

Stage1 (St-1) Conducting literature searches in digital repositories/DB's for most relevant articles.

Stage 2 (St-2) A article's inclusion or removal is based on its title and abstract readings.

Stage 3 (St-3) To determine if an item should be included or not, the introduction and findings must be reviewed.

Stage 4 (St-4) the inclusion and exclusion of data analysis research are based on a full-text review of the research's findings.

In Stage 5 (St-5) most of the original studies that will be included in the SMS study have been vetted and selected for inclusion.

There were 4827 documents collected from the chosen web-based libraries/DB's by imposing inclusion and exclusion criteria following the initial search string

Table 1 Quality evaluation criteria

S.No	Quality evaluation criteria	Options

1	Are the article's outcomes and consequences well-explained?	Yes = 2, No = 0
		Partially = 1

2	Is the study concerned with DB privacy issues?	Yes = 2, No = 0
		Partially = 1

3	What procedures exist to deal with DB security issues?	Yes = 2, No = 0
		Partially = 1

Table 2 Total number of articles per repository

Search string outcomes					
Digital-repository	St-1	St-2	St-3	St-4	St-5
B	2899	61	55	34	23
E	178	39	32	32	13
A	243	26	23	22	16
C	1426	83	77	47	37
F	19	4	1	0	0
D	52	8	7	5	4
Others/Snowballing	10	9	9	8	7
Total	4827	230	204	148	100

iteration (see Sect. 3.3). (Sects. 3.5.1 and 3.5.2, respectively). The tollgate strategy led to a shortlist of 100 publications that were eventually selected for the research. Quality evaluation criteria were used to evaluate the selected articles (Sect. 3.6). Appendix 1 includes a collection of the publications that were ultimately chosen.

3.8 Extracting and synthesizing content

A survey of the articles reviewed is used to obtain the data. In order to address the questions stated in Sect. 3.1, the entire content of every article has been reviewed, and pertinent data extracted. You can find a precise technique for extracting data in the SMS Protocol.

4 Description of key findings

A comprehensive mapping analysis was used throughout this study to determine current state-of-the-art and privacy issues in data modeling, development, and maintenance. Sections 4.1, 4.2, 4.3, 4.4, 4.5 and 4.6 contain the facts of our observations.

4.1 The current state of the art

RQ1 has been addressed using the below sub-questions as a reference (Sects. 4.1.1, 4.1.2, 4.1.3, 4.1.4, 4.1.5 and 4.1.6).

4.1.1 Stages in the building of a protected database

RQ 1.1 focuses on a reliable DB's most frequently studied stages (design, development, and maintenance). As seen in Table 3, the "design" step was mentioned in most publications at a rate of 27%. There is a 25 percent chance that you'll hear about the "developing" stage. The "maintenance" stage was only mentioned in 5 of our SMS research findings.

4.1.2 Well-known sources for the building of reliable DB's

RQ 1.2 is addressed in the second part of this SMS, which concentrates on the location of the papers chosen for this SMS. For venue and provider type analyses, we looked at five repositories, including A, B, C, D, and E. Tables 4 and 5 exhibit the snowballing method, which we refer to as "others." Several of the papers from these collections were presented at conferences, journals, and workshops/symposia, among other venues. As shown in Table 4, 45 out of 100 articles were published through the conference venue. Secondly, we found that, with a rate of 37 out of 100, a large percentage of the publications came from the journal channel. Workshops and symposiums accounted for 18% of the articles presented.

Table 4 lists a total of 100 articles spanning a wide range of topics related to DB privacy. This indicates that scholars have devoted a great deal of attention to this topic. "International Journal of Information Security(IJIS)", "The International Journal on Very Large Data Bases (VLDB)", "Computers and Security (C&S)", "Digital Investigation (DI)", "Journal of Natural Sciences (JNS)" and "Journal of Zhejiang University SCIENCE A (JZUS-A)" were found to be the most popular publications for privacy mechanisms in secure DB designing, as mentioned in Table 5. We also discovered that the "Annual Computer Security Applications Conference(ACSAC)" and the "International Workshop on Digital Watermarking(IWDW)" are the most often referred articles on the issue of our research. Software engineering and other

Table 3 Distinctive stages of a reliable DB design

Stages	Frequency	Percentage
Design	27	27%
Development	25	25%
Maintenance	5	5%

Table 4 A venue-based sampling of articles for the final collection

Repository	Articles from journal	Articles from conference	Articles from workshop or Symposium	Total
B	1	19	3	23
E	13	0	0	13
A	2	9	5	16
C	16	11	10	37
D	0	4	0	4
Others	5	2	0	7
Total	37	45	18	100

related domains can benefit greatly from DB privacy studies.

4.1.3 Research institutions participating in the construction of a reliable DB

The institution of the first researcher was utilized to determine and evaluate the highly ongoing researching institutes in the field of protected DB's. Table 6 shows the findings for RQ 1.3, which reveal that "University of Florida, USA (UOF)" and "CISUC, University of Coimbra, Portugal (UOC)" produced the most research publications on protected DB's (3 percent, each, out of 100). Ben-Gurion University of the Negev (BGU); RMIT University in Melbourne, Australia; YONSEI University in Seoul; TELECOM Bretagne in Brest, France(ENST); Anna University in Chennai, India (AUC); Huazhong University of Science and Technology in Wuhan (HUST); and George Mason University in Fairfax, Virgin Islands(GMU). BGU has presented two publications for each of the selected research.

4.1.4 The most common kind of DB attacks, according to academic research

RQ 1.4 is concerned with identifying the many kinds of DB attacks that have been recorded. Table 7 shows the three types of incidents: internal, external, and both (internal and external). To effectively understand intrusions, we must combine cyber-attacks with breaches by collaborators. Because both internal and external attacks are mentioned in one article, we refer to this as both (internal and external). Our SMS study's "Both (Internal & External)" attacks had a rate of 52, according to the assessment in Table 7. The bulk of the articles in our SMS survey highlighted "External" attacks with a frequency of 35%. In total, 13 papers in our SMS addressed the topic of "internal" attacks.

4.1.5 Database types that have been identified in the literature

To answer RQ 1.5, we must recognize the various DB's discussed in the literature. Seventeen different DB's have been documented in the research based on the data we gathered from the articles we included in our SMS. Table 8

Table 5 Publishing venues containing more than one selected article

Venue	Repository	Type	Frequency
IJIS	C	Journal	3
VLDB	C	Journal	3
ACSAC	A, B	Conference	2
C&S	E	Journal	2
DI	E	Journal	2
IWDW	C	Workshop	2
JNS	C	Journal	2
JZUS-A	C	Journal	2

Table 6 Security DB building research organizations with more than one chosen article

Institution	Rate	Percentage	Country
UOF	3	3%	USA
UOC	3	3%	Portugal
BGU	2	2%	Israel
RMIT	2	2%	Australia
YONSEI	2	2%	South Korea
ENST	2	2%	France
IIT Kharagpur	2	2%	India
AUC	2	2%	India
HUST	2	2%	China
GMU	2	2%	Virginia, USA

Table 7 DB threats that have been described in the literature

Stage	Rate	Percentage
Internal	13	13%
External	35	35%
Both/(internal & external)	52	52%

shows that of the 100 articles in our SMS survey, 24 papers mentioned the term "Web DB." Secondly, we found that "Commercial DB" appeared in 11 of the 100 articles in our SMS analysis. According to SMS, "multilevel DB and distributed DB" was mentioned in ten publications.

4.1.6 Kinds of database management systems (DBMS) presented in the research

Data management systems (DBMS) are examined in RQ 1.6. In this research, 11 distinct DBMS types have been documented based on our SMS data, which was gathered from a selection of studies. Most of the articles in our SMS survey mentioned an "Oracle DB system" with a 31 out of 100 rate, as shown in Table 9. Secondly, "MySQL DB system" was mentioned in most of the publications in our SMS analysis (23 out of 100). Our SMS research found 21 publications that mentioned the term "SQL Server DB system."

Table 8 DB's that have been mentioned in the research

DB type	Rate	Percentage
Web DB's	24	24%
Commercial DB	11	11%
Multilevel DB's	10	10%
Distributed DB	10	10%
Centralized DB	7	7%
Outsourced DB	3	3%
Statistical DB's	3	3%
Dynamic DB	2	2%
Generic DB	2	2%
Automatic DB	2	2%
Open source DB	2	2%
Document DB	1	1%
Real-time DB	1	1%
Time series DB	1	1%
Conventional DB	1	1%
Deductive DB	1	1%
Sensor DB	1	1%

Table 9 Kinds of DBMS as mentioned in the literature

Kinds of DBMS	Rate	Percentage
Oracle	31	31%
MySQL	23	23%
SQL Server	21	21%
Cassandra	7	7%
Mongo DB	6	6%
PostgreSQL	6	6%
Sybase	5	5%
IBM DB2	3	3%
SQLite	2	2%
Informix	2	2%
HBase	1	1%

4.2 Issues in databases

As demonstrated in Table 10 and Fig. 3 our existing research into DB privacy has uncovered 20 issues from a pool of 100 studies (see Appendix 1).

CC #1 Poor authentication system An unauthorized individual gains access to a DB, harvests vital information, and allows the hostile attacker to violate the safety of certified DB's [1, 51].

CC #2 Database intruders We are talking about when we say "threat database attacks" Anonymous queries (anomalous query attack), Harmful queries (query flood attack), and Inferential Attacks (polyinstantiation issue, aggregate problem).

CC #3: Inadequate database protection Best Strategies Specifications Engineering, Architectural, Planning, and Development all suffer from the absence of proper security procedures.

CC #4 Authorized/Malicious User Threats An authorized individual, employee, or administrator may collect or disclose critical data [52].

CC #5 Inadequate access contro Whenever many persons need access to the information, the risk of data fraud and leakage increases. The access should be restricted and regulated [1]

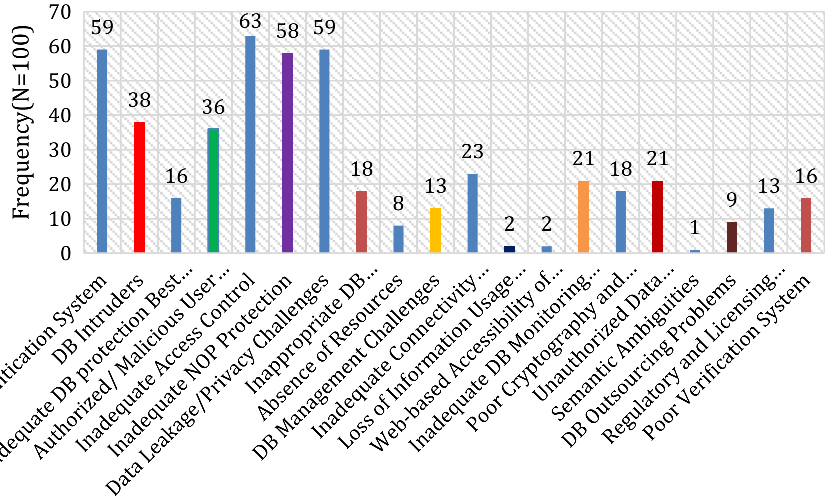
CC #6 Inadequate NOP protection Inadequate NOP Protection is a shortage of network privacy, operating system privacy, and physical safety.

CC #7 Data leakage/privacy challenges Clients of database systems are increasingly concerned about information security. Attacks on disclosed confidential information, including passwords, emails, and private photographs, triggered this issue. Individuals and database systems cannot stop the propagation of data exploitation and destruction once the content has been leaked [53].

Table 10 Challenges in DB security

S. No	Issues	Frequency	Paper ID
1	Poor authentication system	59	1,3,4,5,6,18,19,22,24,25,26,27,28,31,32,33,34,35,36,37,38,39,40,42,43,45,46,47,48,50,51,52,53,54, 59,61,63,64,66,69,73,75,77,79,80,81,82, 83,84,85,86,87,91,92,94,95,96,97,100
2	DB intruders	38	1,2,4,5,6,7,8,9,11,13,15,16,18,21,22,23,26,28,30,34,35,38,42,49,52,53,58,63,64,68,73,75,76,77, 84,89,91,92
3	Inadequate DB protection best strategies	16	1,4,5,10,26,27,32,39,45,47,59,63,80,85,89,93
4	Authorized/malicious user threats	36	4,6,12,15,16,18,19,21,23,26,29,30,31,35,39,43,48,49,51,53,54,55,56,60,61,73,74,75,76,79.81, 82,83,84,85,86,
5	Inadequate access control	63	1,3,4,5,6,7,8,10,13,14,15,16,17,18,20,21,22,23,24,25,26,30,31,33,34,35,36,37,38,39,40,42, 43,44,45,47,48,49, 52,53,54,55,56,59,60,61,69,70,71,73,74,76,77,78, 79,80,83,84,85,87,89,90,91
6	Inadequate NOP protection	58	4,5,6,7,10,13,15,18,19,20,21,22,23,25,26,27,28,29,30,31,35,36,38,39,42,44, 48,49,51,52,53,54,55,56,58,61,62,64,69,70,74,75,76,78,79,80.81,82,83,84,85,88,89,91,93,95,99,100
7	Data leakage/privacy challenges	59	2,5,6,7,8,10,11,12,14,15,16,19,23,26,27,30,32,38,41,42,44,45,46,47,48,49,50,52,53, 54,56,57,61,62,65,66,67,69,70,72,75,76,77,78,79,80,81,82,84,85,87,88,90,91,92,94,95,97,98
8	Inappropriate DB implementation/configuration/maintenance	18	4,6,17,18,20,21,24,26,31,32,35,36,38,39,61,75,85,98
9	Absence of resources	8	10, 41, 61, 62, 70, 91, 96, 98
10	DB management challenges	13	13, 15, 20, 30, 41, 42, 47, 49, 62, 70, 73, 91, 98
11	Inadequate connectivity platforms	23	4, 7, 12, 20, 22, 26, 28, 31, 37, 38, 41, 42, 48, 54, 62, 70, 72, 73, 75, 80, 85, 88, 91
12	Loss of information Usage monitoring	2	26, 41
13	Web-based accessibility of tools for DB attacks	2	26, 42
14	Inadequate DB monitoring strategy	21	5, 7, 16, 19, 20, 21, 22, 26, 28, 30, 31, 42, 51, 56, 70, 84, 86, 82, 91, 94, 99
15	Poor cryptography and anonymization	18	4, 12, 15, 20, 21, 26, 27, 28, 31, 38, 43, 54, 71, 73, 74, 75, 79, 99
16	Unauthorized data alteration/deletion	21	1,26,31,35,39,42,44,54,57,61,66,71,73,74,75,81,82,86,91,95,100
17	Semantic ambiguities	1	40
18	DB outsourcing problems	9	54,57,65,67,69,73,88,96,98
19	Regulatory and licensing challenges	13	1,5,7,12,17,26,34,37,47,70,75,89,94,
20	Poor verification system	16	15,23,24,26,28,31,39,61,63,73,74,75,81,83,84,85

Fig. 3 Issues in DB security



CC #8 Inappropriate database implementation/configuration/maintenance Numerous DB's are improperly setup, formatted, and maintained, among the main reasons for database privacy issues [54].

CC #9 Absence of resources When we talk about a shortage of resources, we are talking about a need of trained employees, a lack of time and budget, a shortage of reliable resources, and an insufficient storage capacity, to name a few things.

CC #10 Database management challenges There are aspects of effectively handling database systems, connectivity, and information at different levels [53].

CC #11 Inadequate connectivity platforms Presently, the majority of customer, user, and third-party conversations are conducted online. The inclusion of an insecure transmission medium was driven by the Internet's opportunity to link DB's [1].

CC #12 Loss of information usage monitoring Several users are unconcerned regarding their communications but may inadvertently send important information to an unauthorized person or untrustworthy servers. Because of a shortage of supervision of data consumption, they are also lost and destroyed [1].

CC #13 Web-based accessibility of tools for database attacks Several tools being used for intrusions are accessible in this globally networked domain, allowing intruders to expose weak spots with minimal expertise of the victim DB architecture [1].

CC #14 Inadequate database monitoring strategy Regulatory risk, discovery, mitigation, and restoration risk are just a few of the dangers posed by a lack of DB auditing [1].

CC #15 Poor cryptography and anonymization No DB privacy plan, regulation, or technology would be sufficient without cryptography, whether the information is traveling over a network or being kept in the DB system [1].

CC #16 Unauthorized data alteration/deletion Any type of unauthorized information alteration or deletion can result in substantial economic losses for an organization or corporation [55].

CC #17 Semantic ambiguities DB issues, including semantic uncertainty, which arises from an absence of semantics or inadequate semantic descriptions, dissemination issues, updating scope constraints, and tuple mistrust, are addressed [56, 57].

CC #18: DB outsourcing problems: Because so many DB's are now being outsourced, there are serious concerns about the *data's accuracy and safety. Clients will have to relinquish management of the information they have outsourced* [58, 59].

CC #19 Regulatory and licensing challenges DB's have many security issues, including policy and licensing concerns. Would the corporation have a consistent and approved policy and licensing from the authorities or organization [1, 60]?

#20 Poor verification system A poor verification system allows an attacker to assume the credentials of a legitimate DB and access its data. The invader has a wide range of options for determining the identification of data. Assuming passwords are easy to remember [1] or using a preset username and password.

4.3 An assessment of database protection issues based on continents

There is much research on various continents in our SMS findings. A comparative analysis of only three continents, i.e., Europe, North America, and Asia, is discussed in this work (See Appendix 2 for more details). We want to find out if these issues are different across continents. We believe that by examining the similarities and distinctions among these problems, we may better prepare ourselves to deal with them on the continent in question. We employed the sequential correlation chi-square test to determine whether there were notable variations among the issues in the three continents listed previously (Martin, 2000). There are many more similarities than distinctions among the issues in the three continents. Poor authentication systems, DB intruders, inadequate DB protection best strategies, and authorized/ malicious user threats are the only major differences found in Table 11. According to our findings, the most prevalent risks in the three continents are "Inadequate Access Control" (65%, 57%, and 64%), "Inadequate NOP Protection" (59%, 57%, and 60%), "Data Leakage/Privacy Challenges" (49%, 60%, 64%), and "Authorized/Malicious Individuals Threats" (40%, 20%, and 52%). It is not uncommon to see "Authorized/Malicious User Threats," "Inadequate Access Control," and "Inadequate NOP Security" across Europe and Asia. Inadequate Connectivity Platforms, Poor Verification Systems, Data Leakage/Privacy Challenges, and Regulatory and Licensing Challenges are some of the problems North American and European clients/users face while creating safe DB's, as shown in Table 11. According to our research, the "Poor Verification System" problem affects the most significant number of customers and users in Asia (78 percent). "Data Leakage/Privacy Challenges" is the most common issue faced by European customers and individuals (60 percent). Many customers in North America face "Inadequate Access Control" and "Data Leakage/Privacy Challenges" concerns, respectively (i.e., 64 percent) (Fig. 4).

4.4 Methodological assessment of database privacy issues

Table 12 shows how we divided the different types of difficulties into three distinct approaches. Table 12 shows the three approaches used: tests, Ordinary literature review OLR, and Other/Mixed Approaches as shown in Fig. 5. Other techniques include writing an experience report, conducting a case study, conducting a survey, and utilizing fuzzy methodologies. When we talk about "many methodologies," we mean that more than one is employed in a single work. Testing is commonly utilized (39 out of 100 times, according to Table 12). The second notable finding in our SMS research is that 31 of the 100 participants used a standard literature review approach. Appendix 2 has further information. Many issues have been revealed by studying the distribution of publications among the three methodologies. Seventeen issues have been detected in relation to OLR, as shown in Table 12. Two of the Seventeen issues have been mentioned in over 50% of the publications. Inadequate Access Control (74%), and Data Leakage/Privacy Challenges (52%), are two of the most often stated problems. Tests face a total of 18 difficulties. Four of these 18 issues have been quoted more than 50% of the time in at least one of the publications. "Data Leakage/Privacy Challenges—64 percent", "Inadequate NOP Protection—62 percent", "Poor Authentication System—56 percent", and "Inadequate Access Control—56 percent" are among the most often stated difficulties. Other/Mixed Approaches publications have highlighted twenty difficulties. Moreover, half of the publications cited 4 of the 20 issues listed. "Poor Authentication System—73%", "Inadequate NOP Protection—63%", "Inadequate Access Control—60%", and "Data Leakage/Privacy Challenges—60%" are among the most frequently stated problems.

Table 12 shows that no SMS approach was employed in any studies (n = 0). These findings prove that our study methodology is innovative in this particular field. We performed the Linear-by-Linear Chi-Square test for the earlier research-mentioned techniques and methodologies to establish whether there was a substantial difference between the challenges. "Poor Authentication System" and "Inappropriate DB implementation/configuration/maintenance" are the only notable variances.

4.5 Years-based study of database privacy issues

A comparison of issues over two time periods, 1990–2010 and 2011–2021, is shown in Table 13 and presented in Fig. 6. More information can be found in Appendix 2. Within the first phase; we found that 18 issues had been highlighted in the research. Four of the 18 issues have been quoted more than 50% in the publications. Inadequate Access Control (70 percent), Poor Authentication System (65 percent), Inadequate NOP Protection (62 percent), and Data Leakage/Privacy Challenges (52 percent) are the most commonly stated vulnerabilities. Between 1990 and 2010, 70 percent of DB's had Inadequate Access Control, indicating that designers failed to effectively control access permission throughout implementationcontrol access permission throughout implementation.

Furthermore, admins in an organization are liable for ensuring that data is adequately protected via access permissions. The "Inadequate Access Control" difficulty has dropped to 58 percent in the second period. The literature

Table 11 Assessment based on the continent

S. No	Issues	Asia (N = 37)		North America (N = 25)		Europe (N = 30)		Chi-square test α = 0.05, df = 1	
		F	%	F	%	F	%	X^2	P
1	**Poor authentication system**	**29**	**78**	**12**	**48**	**14**	**46**	**6.573**	**0.010**
2	**DB intruders**	**10**	**27**	**13**	**52**	**11**	**37**	**3.883**	**0.049**
3	**Inadequate DB protection best strategies**	**9**	**24**	**1**	**4**	**6**	**20**	**3.983**	**0.046**
4	Authorized/malicious user threats	15	40	13	52	6	20	0.416	0.519
5	Inadequate access control	24	65	16	64	17	57	0.022	0.882
6	Inadequate NOP protection	22	59	15	60	17	57	0.000	0.991
7	Data leakage/privacy challenges	18	49	16	64	18	60	1.523	0.217
8	Inappropriate DB Implementation/configuration/maintenance	5	14	4	16	8	27	0.161	0.688
9	Absence of resources	2	5	2	8	2	7	0.164	0.685
10	DB management challenges	2	5	3	12	3	10	0.861	0.354
11	Inadequate connectivity platforms	5	14	4	16	12	40	0.281	0.596
12	Loss of information usage monitoring	1	3	1	4	0	0	0.52	0.819
13	Web-based accessibility of tools for DB attacks	1	3	1	4	0	0	1.151	0.283
14	Inadequate DB monitoring strategy	8	22	5	20	5	17	0.044	0.833
15	Poor cryptography and anonymization	7	19	3	12	8	27	0.284	0.594
16	Unauthorized data alteration/deletion	9	24	5	20	6	20	0.186	0.666
17	Semantic ambiguities	1	3	0	0	0	0	1.151	0.283
18	DB outsourcing problems	4	11	1	4	3	10	0.789	0.374
19	Regulatory and licensing challenges	5	14	2	8	5	17	0.297	0.586
20	Poor verification system	9	24	2	8	5	17	2.750	0.097

Bold indicates significant factors

has revealed 19 problems for the second time period. Four of the 19 obstacles have been referenced in at least half of the publications. “Data Leakage/Privacy Challenges” accounts for 63% of the faults, “Inadequate Access Control” for 58%, “Poor Authentication System” for 55%, and “Inadequate NOP Protection” for 55% of the issues, respectively. We used the Linear-by-Linear Chi-Square analysis and only identified a substantial variation for one problem, “DB Management Challenges, “with a *p*-value of less than.05.

4.6 Evaluation of articles based on their venue

Table 14 displays a breakdown of the various distribution methods. In addition to Journals, Symposiums, Conferences, and Workshops, we have presented our final articles on extracting data via SMS in various other publications venues as well. Journals, Workshops/Symposiums, and conferences have been classified into three categories for easy study. We found that 45 percent of our comprehensive study of articles was presented at conferences, according to Table 14 and Fig. 7. Additionally, 37% of the publications in Table 14 were presented in new journals. For further information, please see Appendix 2 at the ending of the study. Many issues have been discovered as a result of distributing papers via these three channels. According to our findings, 18 issues with journals need to be addressed. Four of the 18 challenges have been referenced in at least half of the publications. “Privacy Issues/Data Leakage—84 percent,” “Inadequate Access Control”—59 percent, “Inadequate NOP Protection”—59 percent,” and “Poor Authentication System—54 percent” are the most often stated difficulties. Conferences face a total of 20 obstacles. Three of these 20 difficulties have been quoted more than 50% of the time in at least one publication. “Poor Authentication System—71 percent,” “Inadequate Access Control—69 percent,” and “Inadequate NOP Protection—62 percent” are the most often stated issues. Workshops/Symposiums face a total of 16 difficulties. Two issues have been mentioned in over half of the publications out of the 16 total. “Data Leakage/Privacy Challenges—61 percent” and “Inadequate Access Control—56 percent” are the most

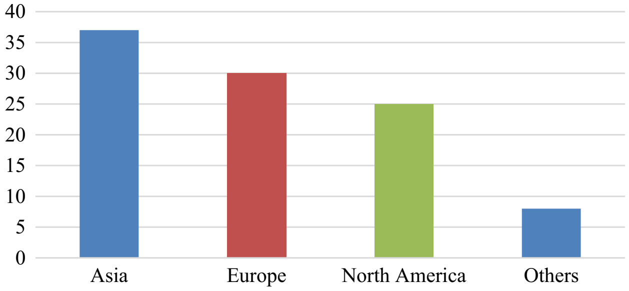


Fig. 4 Distribution depending on continents

commonly reported hurdles. Linear-by-Linear Chi-Squared test has been used to find substantial differences throughout the difficulties. We have found just one big variation between the hurdles "Data Leakage/Privacy Challenges".

4.7 Comparison with existing studies

A wealth of studies have delved into various aspects of database security. Some of these have centered their attention on securing data transmission from server to client, while others have prioritized the construction of secure databases through secure coding practices. The increasing dependence on geographically dispersed information systems for daily operations might augment productivity and efficiency but simultaneously heightens the risk of security violations. Current security measures ensure data transmission protection, yet a comprehensive security strategy must also encompass mechanisms to enforce diverse access control policies. These policies should consider the content sensitivity, data attributes and traits, and other contextual data such as timing.

The consensus in the field is that effective access control systems should integrate data semantics. Moreover, strategies ensuring data integrity and availability must be customized for databases. Consequently, the database security community has developed an array of strategies and procedures over time to safeguard the privacy, integrity, and accessibility of stored data.

Nonetheless, despite these advancements, fresh challenges persist in the database security landscape. Evolving threats, data access "disintermediation," and emerging computing paradigms and applications like grid-based computing and on-demand business have all introduced new security demands and innovative contexts where

Table 12 Methodological based assessment

S. No	Issues	OLR (N = 31)		Experiment (N = 39)		Other/Mixed methods (N = 30)		Chi-square test (Linear-by-linear association) α = 0.05, df = 1	
		F	%	F	%	F	%	X^2	P
1	**Poor authentication system**	**15**	**48**	**22**	**56**	**22**	**73**	**3.866**	**0.049**
2	DB intruders	15	48	15	38	8	27	3.019	0.082
3	Inadequate DB protection best strategies	5	16	5	13	6	20	0.163	0.687
4	Authorized/malicious user threats	10	32	16	41	10	33	0.009	0.924
5	Inadequate access control	23	74	22	56	18	60	1.330	0.249
6	Inadequate NOP protection	15	48	24	62	19	63	1.398	0.237
7	Data leakage/privacy challenges	16	52	25	64	18	60	0.450	0.502
8	**Inappropriate DB implementation/configuration/maintenance**	**11**	**35**	**3**	**8**	**4**	**13**	**5.115**	**0.024**
9	Absence of resources	3	10	2	5	3	10	0.001	0.970
10	DB management challenges	5	16	3	8	5	17	0.002	0.961
11	Inadequate connectivity platforms	6	19	9	23	8	27	0.456	0.500
12	Loss of information usage monitoring	0	0	0	0	2	7	3.379	0.066
13	Web-based accessibility of tools for DB attacks	0	0	1	3	2	7	2.299	0.129
14	Inadequate DB monitoring strategy	9	29	6	15	6	20	0.762	0.383
15	Poor cryptography and anonymization	6	19	5	13	7	23	0.153	0.696
16	Unauthorized data alteration/deletion	6	19	6	15	9	30	1.008	0.315
17	Semantic ambiguities	0	0	0	0	1	3	1.673	0.196
18	DB outsourcing problems	1	3	5	13	3	10	0.866	0.352
19	Regulatory and licensing challenges	5	16	3	8	5	17	0.002	0.961
20	Poor verification system	3	10	8	21	5	17	0.563	0.453

Bold indicates significant factors

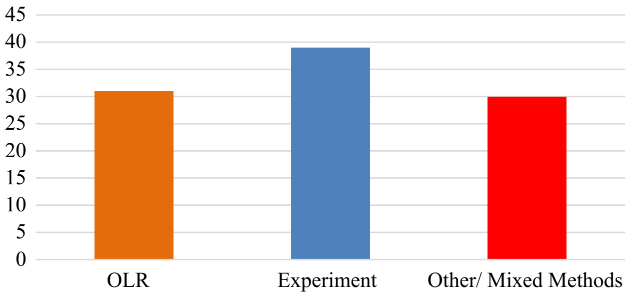


Fig. 5 Methodological-based distribution of papers

existing methodologies can be employed or extended. Despite a multitude of available solutions, raising awareness about existing security breaches is critical for bolstering database security.

In response, we decided to conduct a Systematic Mapping Study (SMS) on secure databases to offer an up-to-date perspective for both database users and developers. We did not find any comprehensive systematic literature review (SLR) or mapping study on this topic to draw comparisons with. However, we believe this research will offer a strategic roadmap for all database stakeholders.

5 Practical implications of research

The practical implications of this research are manifold and impactful. Initially, the results of this SMS will serve as an invaluable resource for DB privacy professionals and users. By leveraging the insights from this study, experts gain an enhanced understanding of DB privacy issues that need addressing. Consequently, they can prioritize their focus on the most significant security challenges. This, in turn, equips DB users with an awareness of their potential privacy risks. Thus, this study benefits consumers by assisting organizations in developing secure DB systems, mindful of the challenges they face (Table 10).

Furthermore, professionals such as DB designers, project managers, and scholars specializing in secure DB design are keen to keep abreast of the latest developments. This research provides DB developers with insights into novel strategies for DB security and the latest advancements in DB technology. Journals such as "VLDB," "Computers & Security," "DI," and "JNS" should be of particular interest to them. Consequently, they would find it beneficial to scrutinize papers available from the "ACSAC" and "IWDW" Conferences and Workshops.

Table 13 Years based evaluation

S. No	Issues	1990–2010 (N = 40)		2011–2021 (N = 60)		Chi-square test (Linear-by-linear association) α = 0.05, df = 1	
		F	%	F	%	X^2	P
1	Poor authentication system	26	65	33	55	0.982	0.322
2	DB intruders	19	48	19	32	2.528	0.112
3	Inadequate DB protection best strategies	7	17	9	15	0.110	0.740
4	Authorized/malicious user threats	15	37	21	35	0.064	0.800
5	Inadequate access control	28	70	35	58	1.387	0.239
6	Inadequate NOP protection	25	62	33	55	0.549	0.459
7	Data leakage/privacy challenges	21	52	38	63	1.153	0.283
8	Inappropriate DB implementation/configuration/maintenance	6	15	12	20	0.402	0.526
9	Absence of resources	1	2	7	12	2.713	0.100
10	**DB management challenges**	**1**	**2**	**12**	**20**	**6.434**	**0.011**
11	Inadequate connectivity platforms	7	17	16	27	1.127	0.288
12	Loss of information usage monitoring	0	0	2	3	1.347	0.246
13	Web-based accessibility of tools for DB attacks	0	0	3	5	2.041	0.153
14	Inadequate DB monitoring strategy	10	25	11	18	0.637	0.425
15	Poor cryptography and anonymization	4	10	14	23	2.862	0.091
16	Unauthorized data alteration/deletion	7	17	14	23	0.487	0.485
17	Semantic ambiguities	1	2	0	0	1.500	0.221
18	DB outsourcing problems	2	5	7	12	1.286	0.256
19	Regulatory and licensing challenges	7	17	6	10	1.182	0.277
20	Poor verification system	5	13	11	18	0.602	0.438

Bold indicates significant factors

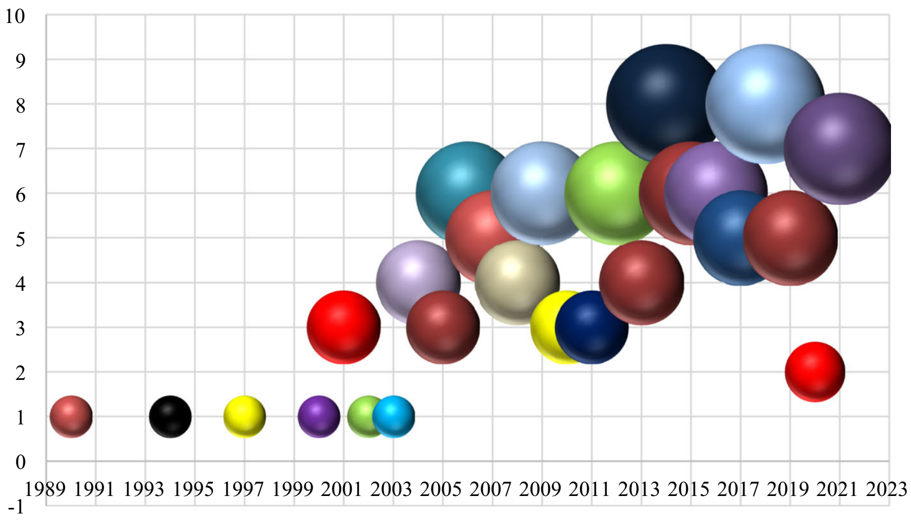


Fig. 6 Year-based distribution of publications

Table 14 Articles venue-based assessment

S. No	Challenges	Journal (N = 37)		Conference (N = 45)		Workshop/ Symposium (N = 18)		Chi-Square Test (Linear-by-Linear Association) α = 0.05, df = 1	
		F	%	F	%	F	%	X^2	P
1	Poor authentication system	20	54	32	71	7	39	0.255	0.613
2	DB intruders	10	27	21	47	7	39	1.456	0.228
3	Inadequate DB protection best strategies	4	11	9	20	3	17	0.597	0.440
4	Authorized/malicious user threats	11	30	18	40	7	39	0.674	0.412
5	Inadequate access control	22	59	31	69	10	56	0.000	0.993
6	Inadequate NOP protection	22	59	28	62	8	44	0.702	0.402
7	**Data leakage/privacy challenges**	**31**	**84**	**17**	**38**	**11**	**61**	**6.153**	**0.013**
8	Inappropriate DB implementation/configuration/maintenance	1	3	15	33	2	11	2.250	0.110
9	Absence of resources	5	14	2	4	1	6	1.610	0.205
10	DB management challenges	7	19	6	13	0	0	3.495	0.062
11	Inadequate connectivity platforms	8	22	12	27	3	17	0.043	0.835
12	Loss of information usage monitoring	0	0	2	4	0	0	0.142	0.706
13	Web-based accessibility of tools for DB attacks	2	5	1	2	0	0	1.354	0.245
14	Inadequate DB monitoring strategy	7	19	11	24	3	17	0.000	0.997
15	Poor cryptography and anonymization	4	11	9	20	5	28	2.250	0.110
16	Unauthorized data alteration/deletion	7	19	9	20	5	28	0.460	0.498
17	Semantic ambiguities	0	0	1	2	0	0	0.070	0.791
18	DB outsourcing problems	5	14	2	4	2	11	0.391	0.532
19	Regulatory and licensing challenges	2	5	9	20	2	11	1.039	0.308
20	Poor verification system	2	5	11	24	3	17	2.339	0.126

Bold indicates significant factors

The aforementioned venues present optimal resources for studying reliable DB development.

These venues, recognized for their focus on secure DB design, encourage scholars to contribute high-quality academic articles. The outcomes of this study will inform experts' decision-making processes, providing guidance on where to invest when developing tools and methodologies for safeguarding DB systems. Lastly, it underscores the need for organizations to provide appropriate training for their customers to tackle critical challenges.

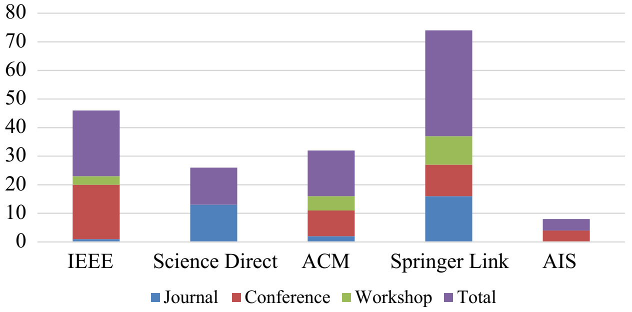


Fig. 7 Venue-based distribution of articles

6 Conclusions and future work

Databases (DBs) serve as repositories for a company's most vital and valuable data, and the increasing popularity of DB systems has been paralleled by a surge in data breach incidents. Therefore, safeguarding DBs necessitates a heightened level of vigilance relative to other systems in the company. As such, a repertoire of comprehensive DB protection measures is needed to uphold the DB system's integrity [1].

Our research employs an SMS approach to explore the issues that various DBs encounter. The derived conclusions are elaborated in the context of the research questions within subsequent subsections. Research Question 1 (RQ1) aims to identify recent DB studies from a software engineering perspective, with the findings detailed in Sect. 4.1 anticipated to inform future research in this realm.

Additionally, we probed the research to answer underlying sub-questions. RQ 1.1 aims to discern the most frequently addressed stages in the literature: planning, developing, and maintaining. In our comprehensive collection of articles, the "design" phase was mentioned with a 27% frequency, as per Table 3. The "development" phase was the second most frequently mentioned, at 25%. The "maintenance" stage was included in 5% of the studies, as shown in Table 3.

RQ 1.2 prompted our selection of five digital repositories as primary locations for article searches. The final pool of articles from these repositories fell into three main publishing categories: Journals, Conferences, and Workshops/Symposiums. Table 4 reveals that 45% of our articles were presented at conferences, while 37% were published in journals. Workshops and symposia contributed 18% of the articles, as per Table 4. Table 5 showcases over 100 articles on DB privacy, indicating a significant commitment of resources to this area. There is a discernible link between the most common venues for secure DB design measures and the venues for secure and robust DB design strategies, as illustrated in Table 5. We also found that "ACSAC" and "IWDW" are the most popular sources of pertinent articles on the topic.

Various institutions continue to contribute significantly to research on secure DBs, which could benefit other areas such as software engineering. The first researcher's affiliations were used to identify the most active research institutions in this area. Table 6 presents the findings for RQ 1.3, showing "UOF" and "UOC" as the most prolific authors of secure DB publications (3% of reviewed publications). "BGU", "YONSEI", "RMIT", "IIT Kharagpur", "ENST", "HUST", "AUC", and "GMU" have each contributed two publications to the selected studies.

RQ 1.4 aimed to identify various types of DB threats, as evidenced in the literature. Table 7 outlines three types of attacks: internal, external, and a combination of both. The category "Both (Internal & External)" accounts for 52% of the discussed studies, according to Table 7. The literature reviewed included 35% of studies examining "External" attacks and 13% discussing "Internal" attacks.

RQ 1.5 sought to identify different DB types addressed in the literature. Seventeen unique DB types were found in the literature, based on data from our SMS-selected publications. As per Table 8, "Web DB" was mentioned in 24% of publications, with "Commercial DB" mentioned in 11%. "Multilevel DB and Distributed DB" were the focus of ten papers in our review.

RQ 1.6 aimed to highlight various types of DB management systems (DBMS) featured in the literature. The findings of our study reveal 11 unique DBMS. Table 9 indicates that "Oracle DB system" was discussed in 31% of publications, while "MySQL DB system" appeared in 23% of articles. "SQL Server" was mentioned in 21 articles, according to our findings.

RQ2 examined DB privacy issues documented in the literature. Applying our predefined inclusion and exclusion criteria, we compiled a finalized set of 100 papers. The data extraction and synthesis process led to a list of 22 challenges, which were narrowed down to 20 upon external reviewer's recommendations, as shown in Table 10. The most common challenges were Data Leakage/Privacy Challenges (63), Poor Authentication System (59), DB Intruders (58), and Authorized/Malicious User Threats (36). These highlighted issues allow companies to evaluate their current security strengths and weaknesses and develop solutions promptly.

If left unaddressed, these identified challenges can potentially inflict harm on the DB system. As such, they must be addressed to enhance DB security. In future work, we aim to identify best practices to alleviate the secure DB challenges outlined in this study.

Appendix 1

See Table 15.

Table 15 Titles of the final selected papers

Paper ID	Title of the final selected papers of SMS
1	Database security
2	Formal query languages for secure relational databases
3	RBAC support in object-oriented role databases
4	The role of cryptography in database security
5	Database security curriculum in infosec program
6	A database security course on a shoestring
7	Fine-grained access control to web databases
8	Privacy leakage in multi-relational databases: a semi supervised learning perspective
9	DROP TABLE textbooks (An argument for SQL injection coverage in database textbooks)
10	Hardening web applications using a least privilege DBMS access model
11	Performance trade-offs on a secure multi-party relational database
12	Securing outsourced database: architecture for protected web resource
13	SeSQLite: security enhanced SQLite (mandatory access control for android databases)
14	Explainable security for relational databases
15	Self-protecting and self-optimizing database systems: implementation and experimental evaluation
16	Teaching database security and auditing
17	A dynamic query-rewriting mechanism for role based access control in databases
18	Database intrusion detection: defending against the insider threat
19	Investigative data warehousing and mining for database security
20	Towards a NoSQL security map
21	A hybrid architecture for database intrusion preventer
22	A practical database intrusion detection system framework
23	Database intrusion detection system using octraplet and machine learning
24	Multilayer access for database protection
25	A flexible database security system using multiple access control policies
26	Database security threats: a survey study
27	Securing database management systems using RAM serial numbers
28	Data masking techniques for NoSQL database security: a systematic review
29	DAIS: a real-time data attack isolation system for commercial database applications
30	Intrusion detection in role administrated database: transaction- based approach
31	Detection of malicious transactions in DBMS
32	Implementing security technique on generic database
33	Design of a new intrusion detection system based on database
34	Security issues in databases
35	Policy-based enforcement of database security configuration through autonomic capabilities
36	Selecting software packages for secure database installations
37	A different approach of intrusion detection and response system for relational databases
38	Detection and database security for a business environment
39	Database intrusion detection by transaction signature
40	Comparative study of poly instantiation models in MLS database
41	A proposal for a reduced client workload model for querying encrypted databases in cloud
42	Analysis for the evaluation and security management of a database in a public organization to mitigate cyber attacks
43	Improvement of ETSFS algorithm for secure database
44	Towards a forensic-aware database solution: using a secured database replication protocol and transaction management for digital investigations
45	Designing secure databases
46	Principles of security and integrity of databases
47	A methodology for integrating access control policies within database development

Table 15 (continued)

Paper ID	Title of the final selected papers of SMS
48	SecureNoSQL: An approach for secure search of encrypted NoSQL databases in the public cloud
49	Machine learning approach to detect intruders in database based on hexplet data structure
50	Discrete algorithms and methods for security of statistical databases related to the work of Mirka Miller
51	Real-time data attack isolation for commercial database applications
52	An access and inference control model for time series databases
53	Anomalous query access detection in RBAC administered databases with random forest and PCA
54	Implementing a database encryption solution, design and implementation issues
55	Database intrusion detection using role and user behavior based risk assessment
56	Carving database storage to detect and trace security breaches
57	Dynamic watermarking-based integrity protection of homomorphically encrypted databases–application to outsourced genetic data
58	Classification agent-based techniques for detecting intrusions in databases
59	A pattern based approach for secure database design
60	Secure deductive databases
61	Multilevel policy based security in distributed database
62	A survey on wireless sensor network databases
63	Prevention guidelines of SQL injection database attacks: an experimental analysis
64	Learning SQL for database intrusion detection using context-sensitive modelling (extended abstract)
65	Encryption techniques for secure database outsourcing
66	Adapted quantization index modulation for database watermarking
67	Verifiable auditing protocol with proxy re-encryption for outsourced databases in cloud
68	Minimizing databases attack surface against SQL injection attacks
69	Secure concurrency control in firm real-time database systems
70	DB-SECaaS: a cloud-based protection system for document-oriented NoSQL databases
71	Data security mechanisms implemented in the database with universal model
72	Secure and efficient anonymization of distributed confidential databases
73	Hierarchical role-based access control with homomorphic encryption for database as a service
74	Process mining and security: visualization in database intrusion detection
75	Integrated intrusion detection in databases
76	Detecting anomalous access patterns in relational databases
77	Combinatorial algorithms and methods for security of statistical databases related to the work of mirka miller
78	Multilevel secure database on security enhanced linux for system high distributed systems
79	Database intrusion detection using sequence alignment
80	A reference model for database security proxy
81	The design and implementation of a self-healing database system
82	Detection of database intrusion using a two-stage fuzzy system
83	Intelligent multi-agent based database hybrid intrusion prevention system
84	A genetic-algorithm based neural network short-term forecasting framework for database intrusion prediction system
85	Server-side database credentials: a security enhancing approach for database access
86	Random forests with weighted voting for anomalous query access detection in relational databases
87	A fine-grained access control model for relational databases
88	Enforcing privacy in cloud databases
89	A comprehensive approach to anomaly detection in relational databases
90	Confidentiality vs integrity in secure databases
91	Security and privacy for multimedia database management systems
92	Privacy leakage in multi-relational databases: a semi-supervised learning perspective
93	Formal analysis on an extended security model for database systems
94	Cloud data security model using modified decoy technique in fog computing for E-healthcare
95	A parallelized database damage assessment approach after cyberattack for healthcare systems

Table 15 (continued)

Paper ID	Title of the final selected papers of SMS
96	A review on autonomous remote security and mobile surveillance using internet of things
97	Securing distributed database using extended blowfish algorithm
98	A review of database services and service providers
99	Analysis of NoSQL database state-of-the-art techniques and their security issues
100	Relational database watermarking techniques: a survey

Appendix 2

See Table 16.

Table 16 Distribution of articles based on Year, methodology, publication channel, and Country/Continent

Paper ID	Year	Authors	Methodology	Publication	Country/ Continent	Venue
1	1990	Teresa F. Lunt, Eduardo B. Fernandez	ORL	Conference	California/ North America	ACM SIGMOD recordSpecial interest group on management of data
2	1994	Marlanne Winslett, Kenneth Smith, Xiaolei Qian	OLR	Journal	Illinois/ North America	Transactions on database systems
3	1997	Raymond K. Wang	Survey + OLR	Conference	Australia	RBAC97: second ACM workshop on role-based access control Fairfax Virginia USA November, 1997 (conference)
4	2004	Ueli Maurer	OLR	Conference	Switzerland/ Europe	International conference on management of data and Symposium on principles database and systems
5	2005	S. Srinivasan, Anup Kumar	OLR	Conference	USA/North America	Information security curriculum development (InfoSecCD) conference
6	2006	Binto George, Anna Valeva	OLR	Conference	USA/North America	ACM SIGCSE Bulletin Special interest group of computer science education
7	2007	Alex Roichman, Ehud Gudes	OLR	Symposium	Israel/Asia	12th ACM Symposium on access control models and technologies
8	2006	Hui Xiong, Michael Steinbach, Vipin Kumar	OLR	Journal	USA/North America	The VLDB journal VLDB
9	2019	Cynthia Taylor, Saheel Sakharkar	OLR	Symposium	USA/North America	ACM technical Symposium on computer science education
10	2018	Stuart Steiner, Daniel Conte de Leon, Ananth A. Jillepalli	OLR + Case study	Symposium	Russia/ Europe	Cyber security Symposium
11	2017	Rogério Pontes, Mário Pinto, Manuel Barbosa	OLR	Symposium	Portugal/ Europe	Symposium on applied computing
12	2015	Kirill Shatilov, Sergey Krendelev, Diana Anisutina, Artem Sumaneev, Evgeny Ogurtsov	OLR	Conference	Russia/ Europe	Central and Eastern European software engineering conference
13	2015	Simone Mutti, Enrico Bacis, Stefano Paraboschi	OLR	Conference	Italy/Europe	ACSAC
14	2014	Gabriel Bender, Łucja Kot, Johannes Gehrke	OLR	Conference	USA/North America	International conference on management of data

Table 16 (continued)

Paper ID	Year	Authors	Methodology	Publication	Country/ Continent	Venue
15	2013	Firas B. Alomari, Daniel A. Menascé	Experiment	Conference	Virginia/ North America	Cloud and autonomic computing conference
16	2009	Li Yang	OLR	Symposium	USA/North America	Technical Symposium on computer science
17	2008	Jay Jarman, James A. McCart, Donald Berndt, Jay Ligatti	OLR	Conference	USA/North America	Americas conference on information systems
18	2012	Kevin A. Barton, Carol J. Jeffries-Horner	OLR	Conference	USA/North America	Eighteenth Americas conference on information systems
19	2006	Areej Yassin, Donald Berndt, Monica Chiarini	OLR	Conference	USA/North America	Americas conference on information systems
20	2018	Wilhelm Zugaj, Anita Stefanie Beichler	OLR	Conference	Austria/ Europe	International conference on information systems development
21	2014	Yousef Ali Albakoush, Roslan Ismail, Asmidar Abu Bakar	Experiment	Conference	Malaysia/ Asia	International conference on information technology and multimedia
22	2009	Yawei Zhang and Xiaojun Ye, Feng Xie and Yong Peng	Experiment	Conference	China/Asia	International conference on computer and information technology
23	2018	Souparnika Jayaprakash, Kamalanathan Kandasamy	Experiment	Conference	India/Asia	International conference on inventive communication and computational technologies
24	2018	Olesia Voitovych, Leonid Kupershtein, Vitalii Lukichov, Ivan Mikityuk	OLR	Conference	Ukraine/ Europe	International scientific-practical conference problems of infocommunications. Science and technology
25	2003	Min-A Jeong, Jung-Ja Kim, and Yonggwan Won	OLR	Conference	Korea/Asia	International conference on parallel and distributed computing, applications and technologies
26	2013	Nedhal A. Al-Sayid, Dana Aldlaeen	Survey	Conference	Jordan/Asia	International conference on computer science and information technology
27	2019	Sapan Noori Azeez, Serkan Varol	OLR + Experiment	Symposium	Turkey/ Asia/ Europe	International Symposium on digital forensics and security
28	2017	Alfredo Cuzzocrea, Hossain Shahriar	OLR + Experiment	Conference	Italy/Europe	International conference on big data
29	2001	Peng Liu	Experiment	Conference	USA/North America	ACSAC
30	2013	Saad M. Darwish, Shawkat K. Guirguis, Mahmoud M. Ghozlan	OLR	Conference	Egypt/ Africa/ Asia	International conference on computer engineering & systems
31	2005	Marco Vieira, Henrique Madeira	OLR	Symposium	Portugal/ Europe	Pacific rim international Symposium on dependable computing
32	2015	Gaurav Dubey, Vikram Khurana, Shelly Sachdeva	OLR	Conference	India/Asia	International conference on contemporary computing
33	2009	Gongxing Wu, Yimin Huang	OLR + Experiment	Conference	China/Asia	International conference on signal processing systems
34	2009	Sohail IMRAN, Dr. Irfan Hyder	OLR + Experiment	Conference	Pakistan/ Asia	International conference on future information technology and management engineering
35	2008	Ghassan "Gus" Jabbour, Daniel A. Menasć	OLR	Conference	USA/North America	International conference on autonomic and autonomous systems

Table 16 (continued)

Paper ID	Year	Authors	Methodology	Publication	Country/ Continent	Venue
36	2011	Afonso Araújo Neto, Marco Vieira	Survey + Experiment	Conference	Portugal/ Europe	International conference on availability, reliability and security
37	2013	Jitendra Parmar, Pranita jain	Experiment	Conference	India/Asia	International conference on green computing, communication and conservation of energy
38	2012	Traian Popeea, Anca Constantinescu, Laura Gheorghe, Nicolae Tăpus	Experiment	Conference	Romania/ Europe	International conference on intelligent networking and collaborative systems
39	2012	Yagnik A Rathod, Prof. M.B. Chaudhari, Prof. G.B. Jethava	OLR + Experiment	Conference	India/Asia	International conference on computing, communication and networking technologies
40	2010	Ahmed I. Sallam, Sayed M. Elrabie, and Osama S. Faragallah	Survey + Experiment	Conference	Egypt/ Africa/ Asia	International computer engineering conference
41	2019	Víctor Alexis Fuentes Tello, Brajendra Panda	OLR + Experiment	Conference	Panama/ North America	Latin American computing conference
42	2020	Segundo M. Toapanta, Omar A. Escalante, Luis E. Mafla, and Rocío M. Arellano	OLR + Experiment	Journal	Ecuador/ South America	Digital object identifier
43	2016	Prathyusha Uduthalapally, Bing Zhou	OLR + Experiment	Symposium	USA/North America	International Symposium on digital forensic and security
44	2014	Peter Frühwirt, Peter Kieseberg, Katharina Krombholz, Edgar Weippl	OLR + Experiment	Journal	Austria/ Europe	Journal DI
45	2004	Eduardo Fernández-Medina, Mario Piattini	Case study	Journal	Spain/ Europe	Journal information and software technology
46	2014	Serban Mariuta	Experiment	Journal	Romania/ Europe	Journal procedia economics and finance
47	2012	Jenny Abramov, Omer Anson, Michal Dahan, Peretz Shoval, Arnon Sturm	Survey	Journal	Israel/Asia	Journal computers & security
48	2016	Mohammad Ahadian, Frank Plochan, Zak Roessler, Dan C. Marinescu	Experiment	Journal	USA/North America	International journal of information management
49	2016	Saad M. Darwish	OLR + Experiment	Journal	Egypt/ Africa/ Asia	Journal of electrical systems and information technology
50	2018	Andrei Kelarev, Joe Ryan, Leanne Rylands, Jennifer Seberry, Xun Yi	Survey	Journal	Australia	Journal of discrete algorithms
51	2006	Peng Liu, Hai Wang, Lunquan Li	OLR + Experiment	Journal	USA/North America	Journal of network and computer applications
52	2019	Amir Noury, Morteza Amini	Experiment	Journal	Iran/Asia	Journal of future generation computer systems
53	2016	Charissa Ann Ronao, Sung-Bae Cho	Experiment	Journal	Korea/Asia	Journal information sciences
54	2014	Erez Shmueli, Ronen Vaisenberg, Ehud Gudes, Yuval Elovici	Survey + Experiment	Journal	USA/North America	Journal computers & security
55	2020	Indu Singh, Narendra Kumar, Srinivasa K.G., Tript Sharma, Vaibhav Kumar, Siddharth Singhal	Experiment	Journal	India/Asia	Journal of information security and applications

Table 16 (continued)

Paper ID	Year	Authors	Methodology	Publication	Country/ Continent	Venue
56	2017	James Wagner, Alexander Rasin, Boris Glavic, Karen Heart, Jacob Furst, Lucas Bressan, Jonathan Grier	Experiment	Journal	USA/North America	Journal DI
57	2018	David Niyitegeka, Gouenou Coatrieux, Reda Bellafqira, Emmanuelle Genin, and Javier Franco-Contreras	Experiment	Workshop	France/ Europe	IWDW
58	2008	Cristian Pinzón, Yanira De Paz, and Rosa Cano	Survey	Workshop	Panama/ North America	International workshop on hybrid artificial intelligence systems
59	2011	Arnon Sturm, and Peretz Shoval, Jenny Abramov	OLR	Workshop	Israel/Asia	International conference on advanced information systems engineering
60	2001	Steve Barker	OLR	Symposium	UK/Europe	International Symposium on practical aspects of declarative languages
61	2011	Neera Batra and Manpreet Singh	OLR	Conference	India/Asia	International conference on advances in computing and communications
62	2019	Abderrahmen Belfkih, Claude Duvallet, Bruno Sadeg	Survey + OLR	Journal	France/ Europe	Journal of mobile communication, computation and information
63	2015	Vijaylaxmi Bittal and Soumi Banerjee	Experiment	Conference	India/Asia	Conference emerging research in computing, information, communication and applications
64	2009	Christian Bockermann, MartinApel, Michael Meier	Experiment	Conference	Germany/ Europe	International conference on detection of intrusions and malware, and vulnerability assessment
65	2007	Sergei Evdokimov and Oliver Ğunther	Experiment	Symposium	Germany/ Europe	European Symposium on research in computer security
66	2014	Javier Franco-Contreras, Gouenou Coatrieux, Nora Cuppens-Boulahia, Fŕeď eric Cuppens, and Christian Roux	Experiment	Workshop	France/ Europe	IWDW
67	2018	GAO Ziyuan, WANG Baocang, LIU Hequn, LU Ke, ZHAN Yu	Experiment	Journal	China/Asia	JNS
68	2015	Dimitris Geneiatakis	Experiment	Conference	Greece/ Europe	International conference on information and communications security
69	2000	BINTO GEORGE, JAYANT R. HARITSA	OLR + Experiment	Journal	India/Asia	Security of data and transaction processing
70	2016	Yumna Ghazi, Rahat Masood, Abid Rauf, Muhammad Awais Shibli, and Osman Hassan	Experiment	Journal	Australia	IJIS
71	2014	V. M. Grachev, V.I.Esin, N. G. Polukhina, and S. G. Rassomakhin	OLR	Journal	Ukraine/ Europe	Bulletin of the lebedev physics institute
72	2014	Javier Herranz, Jordi Nin	Experiment	Journal	Spain/ Europe	IJIS
73	2016	Kamlesh Kumar Hingwe and S. Mary Saira Bhanu	Experiment	Conference	India/Asia	International conference on ICT for sustainable development
74	2012	Viet H. Huynh and An N.T. Le	OLR + Experiment	Workshop	Viet Nam/ Asia	Pacific-Asia workshop on intelligence and security informatics

Table 16 (continued)

Paper ID	Year	Authors	Methodology	Publication	Country/ Continent	Venue
75	2007	José Fonseca, Marco Vieira, and Henrique Madeira	Experience Report	Symposium	Portugal/ Europe	Symposium on dependable computing
76	2007	Ashish Kamra, Evimaria Terzi, Elisa Bertino	Experiment	Journal	USA/North America	The VLDB journal VLDB
77	2018	Andrei Kelarev, Jennifer Seberry, Leanne Rylands, and Xun Yi	Survey	Workshop	Australia	International workshop on combinatorial algorithms
78	2012	Haklin Kimm and Norkee Sherpa	Experiment	Workshop	USA/North America	International workshop on information security applications
79	2010	Amlan Kundu, Shamik Sural, A. K. Majumdar	Experiment	Journal	India/Asia	IJIS
80	2002	CM Lima, YANG Xiao-h, DONG Jin-xiang	Experiment	Journal	China/Asia	Journal of Zhejiang University
81	2004	PENG LIU, JIWU JING	Experiment	Journal	USA/ North America	Journal of intelligent information systems
82	2009	Suvasini Panigrahi and Shamik Sural	Fuzzy Methods	Conference	India/Asia	International conference on information security
83	2004	P. Ramasubramanian and A. Kannan	Experiment	Conference	India/Asia	Conference on advances in databases and information systems
84	2006	P. Ramasubramanian, A. Kannan	Experiment	Journal	India/Asia	Journal soft computing
85	2017	Diogo Domingues Regateiro, Óscar Mortágua Pereira, and Rui L. Aguiar	Experiment	Conference	Portugal/ Europe	International conference on data management technologies and applications
86	2015	Charissa Ann Ronao and Sung-Bae Cho	Experiment	Conference	South Korea/ Asia	International conference on artificial intelligence and soft computing
87	2010	Jie SHI, Hong ZHU	Experiment	Journal	China/Asia	Journal of Zhejiang University
88	2017	Somayeh Sobati Moghadam, J́er̂ome Darmont, and Ǵerald Gavin	Survey	Conference	France/ Europe	International conference on big data analytics and knowledge discovery
89	2005	Adrian Spalk and Jan Lehnhardt	Experiment	Conference	Germany/ Europe	Conference on data and applications security
90	2001	Adrian Spalka and Armin B. Cremers	OLR	Journal	Germany/ Europe	Part of the IFIP international federation for information processing book series
91	2007	Bhavani Thuraisingham	OLR	Journal	USA/North America	Journal of multimedia tools and applications
92	2006	Hui Xiong, Michael Steinbach, Vipin Kumar	Experiment	Journal	USA/North America	The VLDB journal VLDB
93	2008	ZHU Hong, ZHU Yi, LI Chenyang, SHI Jie, FU Ge, WANG Yuanzhen	Experiment	Journal	China/Asia	JNS
94	2021	Dr. P. Maragathavalli, S. Atchaya, N. Kaliyaperumal, and S. Saranya	OLR + Experiment	Conference	India/Asia	International conference on frontiers in engineering science and technology
95	2021	Sanaa Kaddoura, Ramzi A. Haraty, Karam Al Kontar, and Omar Alfandi	Experiment	Journal	UAE/Asia	Journal of future internet
96	2021	Manoj Diwakar, Kanika Sharma, Ravi Dhaundiyal, Sheetal Bawane, Kapil Joshi, Prabhishek Singh	Experiment	Journal	India/Asia	Journal of physics
97	2021	Sangeetha Radhakrishnan, Dr. A.Akila	Experiment	Journal	India/Asia	International journal of modern agriculture

Table 16 (continued)

Paper ID	Year	Authors	Methodology	Publication	Country/ Continent	Venue
98	2021	Ukpe Kufre Christopher, Asagba, Prince Oghenekaro	OLR	Journal	Nigeria/ Africa	International journal of computer techniques
99	2021	Harpreet Kaur	OLR	Journal	India/Asia	Journal of computer and mathematics education
100	2021	Asmaa Alqassab, Mafaz Alanezi	Survey	Conference	Iraq/Asia	International conference for pure and applied sciences

Acknowledgements The authors would like to acknowledge the support provided by the Deanship of Scientific Research via project number DF201007 at King Fahd University of Petroleum and Minerals, Saudi Arabia.

Funding Open Access funding provided by University of Oulu including Oulu University Hospital.

References

1. Al-Sayid, N. A. & Aldlaeen, D. (2013). Database security threats: A survey study. In *2013 5th international conference on computer science and information technology*, pp. 60–64.
2. Humayun, M., Jhanjhi, N., Almufareh, M. F., & Khalil, M. I. (2022). Security threat and vulnerability assessment and measurement in secure software development. *Computers, Materials and Continua, 71*, 5039–5059.
3. Afzal, W., Torkar, R., & Feldt, R. (2009). A systematic review of search-based testing for non-functional system properties. *Information and Software Technology, 51*, 957–976.
4. Humayun, M., Jhanjhi, N. Z., & Almotilag, A. (2022). Real-time security health and privacy monitoring for Saudi highways using cutting-edge technologies. *Applied Sciences, 12*, 2177.
5. Toapanta, S. M., Quimis, O. A. E., Gallegos, L. E. M., & Arellano, M. R. M. (2020). Analysis for the evaluation and security management of a database in a public organization to mitigate cyber attacks. *IEEE Access, 8*, 169367–169384.
6. Almufareh, M. F., & Humayun, M. (2023). Improving the safety and security of software systems by mediating SAP verification. *Applied Sciences, 13*, 647.
7. Fernández-Medina, E., & Piattini, M. (2005). Designing secure databases. *Information and Software Technology, 47*, 463–477.
8. Brahma, A. & Panigrahi, S. (2022). Application of soft computing techniques in database intrusion detection. In *Intelligent technologies: concepts, applications, and future directions*, Springer, pp. 201–221.
9. Chakraborty, M. S. (2022). Database security threats and how to mitigate them. In *Empowering Smart Future Through Scientific Development and Technology Conference*, USA. https://doi.org/10.3390/mol2net-08-12642
10. Singh, V., & Yadav, V. (2021). Survey of blockchain applications in database security. In *Advances in distributed computing and machine learning: proceedings of ICADCML 2020*, pp. 147–154.
11. Pevnev, V. & Kapchynskyi, S. (2018). Database security: threats and preventive measures. *Сучасні інформаційні системи*, pp. 69–72.
12. Alisawi, W. C., Hussain, A. A. A., & Alawsi, W. A. (2019). Estimate new model of system management for database security. *Indonesian Journal of Electrical Engineering and Computer Science, 14*, 1391–1394.
13. Agboola, R. B., Iro, Z. S., Awwalu, J. & Said, I. N. (2022). Database security framework design using tokenization. *Dutse Journal of Pure and Applied Sciences, 8*, 16–26.
14. Nagamani, C., & Chittineni, S. (2022). Network database security with intellectual access supervision using outlier detection techniques. *International Journal of Advanced Intelligence Paradigms, 22*, 348–361.
15. Moghadam, S. S., Darmont, J. & Gavin G. (2017). Enforcing privacy in cloud databases. In *International conference on big data analytics and knowledge discovery*, pp. 53–73.
16. Kelarev, A., Seberry, J., Rylands, L. & Yi X. (2017). Combinatorial algorithms and methods for security of statistical databases related to the work of Mirka Miller. In *International workshop on combinatorial algorithms*, pp. 383–394.
17. Mai, X., Zhang, Y., Zhang, T. & Li, M. (2023). Security protection method of power system database based on cloud platform. In *International conference on statistics, data science, and computational intelligence (CSDSCI 2022)*, pp. 234–239.
18. Ibrahim, S., Zengin, A., Hizal, S., Suaib Akhter, A. & Altunkaya, C. (2023). A novel data encryption algorithm to ensure database security. *Acta Infologica, 7*(1), 1–16.
19. Abdulameer, S. A. (2023). A cryptosystem for database security based on RC4 algorithm. *Journal of Qadisiyah for Computer Science and Mathematics., 15*, 189–196.
20. Jayaprakash, S. & Kandasamy, K. (2018). Database intrusion detection system using octraplet and machine learning. In *2018 second international conference on inventive communication and computational technologies (ICICCT)*, pp. 1413–1416.
21. Singh, I., Kumar, N., Srinivasa, K., Sharma, T., Kumar, V., & Singhal, S. (2020). Database intrusion detection using role and

user behavior based risk assessment. *Journal of Information Security and Applications, 55*, 102654.

22. Humayun, M., & Jhanjhi, N. (2019). Exploring the relationship between GSD, knowledge management, trust and collaboration. *Journal of Engineering Science and Technology, 14*, 820–843.
23. Popeea, T., Constantinescu, A., Gheorghe, L. & Tapus, N. (2012). Inference detection and database security for a business environment. In *2012 fourth international conference on intelligent networking and collaborative systems*, pp. 612–617.
24. Yang, P., Xiong, N., & Ren, J. (2020). Data security and privacy protection for cloud storage: A survey. *IEEE Access, 8*, 131723–131740.
25. Uduthalapally, P. & Zhou, B. (2016). Improvement of ETSFS algorithm for secure database. In *2016 4th international symposium on digital forensic and security (ISDFS)*, pp. 63–67.
26. Christopher, U. K. & Asagba, P. O. (2021). A review of database services and service providers. *International Journal of Computer Techniques, 8*.
27. Jain, S., & Chawla, D. (2020). A relative study on different database security threats and their security techniques. *International Journal of Innovative Science and Research Technology, 5*, 794–799.
28. Lawal, M., Sultan, A. B. M., & Shakiru, A. O. (2016). Systematic literature review on SQL injection attack. *International Journal of Soft Computing, 11*, 26–35.
29. Bria, R., Retnowardhani, A., & Utama, D. N. (2018). Five stages of database forensic analysis: A systematic literature review. In *2018 international conference on information management and technology (ICIMTech)*, IEEE, pp. 246–250
30. Cuzzocrea, A. & Shahriar, H. (2017). Data masking techniques for NoSQL database security: A systematic review. in *2017 IEEE international conference on big data (Big Data)*, pp. 4467–4473.
31. Soomro, Z. A., Shah, M. H., & Ahmed, J. (2016). Information security management needs more holistic approach: A literature review. *International Journal of Information Management, 36*, 215–225.
32. Márquez, G., Astudillo, H., & Taramasco, C. (2020). Security in telehealth systems from a software engineering viewpoint: A systematic mapping study. *IEEE Access, 8*, 10933–10950.
33. Mousa, A., Karabatak, M. & Mustafa, T. (2020). Database security threats and challenges. In *2020 8th international symposium on digital forensics and security (ISDFS)*, pp. 1–5.
34. Petersen, K., Vakkalanka, S., & Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology, 64*, 1–18.
35. Staples, M., & Niazi, M. (2008). Systematic review of organizational motivations for adopting CMM-based SPI. *Information and software technology, 50*, 605–620.
36. Humayun, M., Niazi, M., Almufareh, M. F., Jhanjhi, N., Mahmood, S., & Alshayeb, M. (2022). Software-as-a-Service security challenges and best practices: A multivocal literature review. *Applied Sciences, 12*, 3953.
37. Humayun, M., Niazi, M., Jhanjhi, N., Alshayeb, M., & Mahmood, S. (2020). Cyber security threats and vulnerabilities: A systematic mapping study. *Arabian Journal for Science and Engineering, 45*, 3171–3189.
38. Khan, A. A., Keung, J., Niazi, M., Hussain, S. & Zhang, H. (2017). Systematic literature reviews of software process improvement: A tertiary study. In*Systems, software and services process improvement: 24th european conference, EuroSPI 2017, Ostrava, Czech Republic, September 6–8, 2017, Proceedings 24*, pp. 177–190.
39. Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering–a systematic literature review. *Information and software technology, 51*, 7–15.
40. Inayat, I., Salim, S. S., Marczak, S., Daneva, M., & Shamshirband, S. (2015). A systematic literature review on agile requirements engineering practices and challenges. *Computers in human behavior, 51*, 915–929.
41. Taylor, P. J., Dargahi, T., Dehghantanha, A., Parizi, R. M., & Choo, K.-K.R. (2020). A systematic literature review of blockchain cyber security. *Digital Communications and Networks, 6*, 147–156.
42. Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. In *Proceedings of the 18th international conference on evaluation and assessment in software engineering*, pp. 1–10.
43. Sauerwein, C., Gander, M., Felderer, M. & Breu, R. (2016). A systematic literature review of crowdsourcing-based research in information security. In *2016 IEEE symposium on service-oriented system engineering (SOSE)*, pp. 364–371.
44. Mourão, E., Pimentel, J. F., Murta, L., Kalinowski, M., Mendes, E., & Wohlin, C. (2020). On the performance of hybrid search strategies for systematic literature reviews in software engineering. *Information and Software Technology, 123*, 106294.
45. Dissanayake, N., Jayatilaka, A., Zahedi, M. & Babar, M. A. (2020). Software security patch management–a systematic literature review of challenges, approaches, tools and practices. *arXiv preprint* arXiv:2012.00544.
46. Hulshof, M. & Daneva, M. (2021). Benefits and challenges in information security certification–a systematic literature review. In *International symposium on business modeling and software design*, pp. 154–169.
47. Mendes, E., Wohlin, C., Felizardo, K., & Kalinowski, M. (2020). When to update systematic literature reviews in software engineering. *Journal of Systems and Software, 167*, 110607.
48. Ahmad, A., Khan, S. U., Khan, H. U., Khan, G. M., & Ilyas, M. (2021). Challenges and practices identification via a systematic literature review in the adoption of green cloud computing: Client's side approach. *IEEE Access, 9*, 81828.
49. Mohammed, N. M., Niazi, M., Alshayeb, M., & Mahmood, S. (2017). Exploring software security approaches in software development lifecycle: A systematic mapping study. *Computer Standards & Interfaces, 50*, 107–115.
50. Khan, A. A., Keung, J., Niazi, M., Hussain, S., & Ahmad, A. (2017). Systematic literature review and empirical investigation of barriers to process improvement in global software development: Client–vendor perspective. *Information and Software Technology, 87*, 180–205.
51. Shmueli, E., Vaisenberg, R., Gudes, E., & Elovici, Y. (2014). Implementing a database encryption solution, design and implementation issues. *Computers & security, 44*, 33–50.
52. Shatilov, K., Krendelev, S., Anisutina, D., Sumaneev, A. & Ogurtsov, E. (2015). Securing outsourced database: architecture for protected web resource. In *Proceedings of the 11th central & Eastern European software engineering conference in Russia*, pp. 1–7.
53. R. Pontes, M. Pinto, M. Barbosa, R. Vilaça, M. Matos, and R. Oliveira, "Performance trade-offs on a secure multi-party relational database," in *Proceedings of the Symposium on Applied Computing*, 2017, pp. 456–461.
54. George, B., & Valeva, A. (2006). A database security course on a shoestring. *ACM SIGCSE Bulletin, 38*, 7–11.
55. Ronao, C. A. & Cho, S.-B. (2015). Random forests with weighted voting for anomalous query access detection in relational databases. In *International conference on artificial intelligence and soft computing*, pp. 36–48.
56. Sallam, A. I., Elrabie, S. M. & Faragallah, O. S. (2010). Comparative study of polyinstantiation models in MLS database. In

2010 international computer engineering conference (ICENCO), pp. 158–165.

57. Dragos, V. (2021). Semantic frameworks to enhance situation awareness for defence and security applications (Doctoral dissertation, Université de Paris).
58. Niyitegeka, D., Coatrieux, G., Bellafqira, R., Genin, E. & Franco-Contreras, J. (2018). Dynamic watermarking-based integrity protection of homomorphically encrypted databases–application to outsourced genetic data. In *International workshop on digital watermarking*, pp. 151–166.
59. Nadim, M., Latif, R. M. A., Hussain, K., Jhanjhi, N., Masud, M., Alyahyan, S. Y., et al. (2021). A framework for software customization in global software development (GSD). *Turkish Online Journal of Qualitative Inquiry, 12*, 3331–3364.
60. Maragathavalli, P., Atchaya, S., Kaliyaperumal, N. & Saranya, S. (2021). Cloud data security model using modified decoy technique in fog computing for E-healthcare. In *IOP conference series: Materials science and engineering*, p. 012044.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Asif Iqbal Received a Master of Philosophy degree in computer science from University of Malakand, Khyber Pakhtoonkhwa, Pakistan. He completed his M.Phil. degree under the supervision of Dr. Siffat Ullah Khan. He is a member of the Software Engineering Research Group University of Malakand. Currently, he is serving as a government teacher in the Elementary and Secondary Education Department, Khyber Pakhtoonkhwa, Pakistan. His area of interest includes Database security, requirements engineering, system analysis, and Cybersecurity.

Siffat Ullah Khan received a Ph.D. degree in computer science from Keele University, U.K., in 2011. He was the Head of the Department of Software Engineering, University of Malakand, Pakistan, for three years, whereas he also served as the Chairman of the Department of Computer Science and IT and is currently working as an Associate Professor in Computer Science at the University of Malakand. He is also the Founder and the Leader of the Software Engineering Research Group, University of Malakand. He has successfully supervised 19 M.Phil. and six Ph.D. scholars. He has authored over 100 articles so far in well-reputed international conferences and journals. His research interests include cybersecurity, software outsourcing, empirical software engineering, agile software development, systematic literature review, software metrics, cloud computing, requirements engineering, and green computing/IT. He received the Gold Medal (Dr. M. N. Azam Prize 2015) from the Pakistan Academy of Sciences in recognition of his research achievements in the field of computer (software)

Mahmood Niazi is a Professor of Software Engineering at the Information and Computer Science Department, King Fahd University of Petroleum and Minerals Saudi Arabia. He has received the MPhil degree from the University of Manchester, U.K., and the Ph.D. degree from the University of Technology Sydney, Australia. He has spent more than a decade with leading technology firms and universities as a Process Analyst, a Senior Systems Analyst, a Project Manager, and a Professor. He has participated in and managed several software development projects. Dr. Niazi is an active researcher in the field of empirical software engineering. Dr. Niazi has published over 100 articles He is interested in developing sustainable processes in order to develop systems, which are reliable, secure, and fulfill customer needs. His research interests are evidence-based software engineering, requirements engineering, sustainable, reliable, and secure software engineering processes, global and distributed software engineering, software process improvement, and software engineering project management. Previously Dr. Niazi worked for Keele University UK, National ICT Australia, University of Technology Sydney Australia, University of Sydney Australia, and the University of Manchester UK.

Mamoona Humayun has completed her Ph.D. in Computer Sciences from Harbin Institute of Technology, China. She has 15 plus years of teaching and administrative experience internationally. She has extensive background of teaching, research supervision and administrative work. She has experience in teaching advanced era technological courses besides other undergraduate and postgraduate courses, graduation projects and thesis supervisions. Dr. Mamoona Humayun is the guest Editor and reviewer for several reputable journals and conferences around the globe. She has authored several research papers, supervised a great number of postgraduate students, and external thesis examiner to her credit. She has strong analytical, problem solving, interpersonal and communication skills. Her areas of interest include Software Engineering, Cyber Security, Wireless Sensor Network (WSN), Internet of Things (IoT), big data, Requirement Engineering, Global Software Development and Knowledge Management.

Najm Us Sama received the bachelor's degree in computer science from the University of Malakand, Pakistan, in 2007, the master's degree in computer science from the University of Peshawar, Pakistan, in 2009, and the Ph.D. degree in information technology from the University Malaysia Sarawak (UNIMAS), Malaysia, in 2019. She is currently a Teacher with the Faculty of Computer Science and Information Technology (FCSIT), University Malaysia Sarawak, Kota Samarahan, Malaysia. Her research interests include

MANET, image processing, and deployment, coverage, and energy efficient utilization in wireless sensor networks.

Arif Ali Khan received a Ph.D. degree in software engineering from the City University of Hong Kong, Hong Kong. He is currently an Assistant Professor with the M3S Empirical Software Engineering Research Unit, University of Oulu, Finland. He has participated in and managed several empirical software engineering-related research projects. He has expertise in software process improvement, quantum software engineering, microservices architecture, artificial intelligence (AI) ethics, agile software development, DevOps, global software development, multicriteria decision analysis, soft computing, and evidence-based software engineering.

Aakash Ahmad received the Ph.D. degree in software engineering from Dublin City University, Dublin, Ireland (funded by Lero—the Irish Software Engineering Research Center, Ireland). He is currently an Assistant Professor of computing and software engineering with the School of Computing and Communications, Lancaster University Leipzig, Leipzig, Germany. His research interests are in the area of software and service engineering for quantum computing systems.