On Understanding Context Modelling for Adaptive Authentication Systems

Anne Bumiller, Stéphanie Challita, Benoit Combemale, Olivier Barais, Nicolas Aillery, Gael Le Lan

On Understanding Context Modelling for Adaptive Authentication Systems

ANNE BUMILLER, Orange Labs, University of Rennes 1 / IRISA / INRIA, France
STÉPHANIE CHALLITA, University of Rennes 1 / IRISA / INRIA, France
BENOIT COMBEMALE, University of Rennes 1 / IRISA / INRIA, France
OLIVIER BARAIS, University of Rennes 1 / IRISA / INRIA, France
NICOLAS AILLERY, Orange Labs, France
GAEL LE LAN, Orange Labs, France

In many situations, it is of interest for authentication systems to adapt to context (*e.g.,* when the user's behavior differs from the previous behavior). Hence, representing the context with appropriate and well-designed models is crucial. We provide a comprehensive overview and analysis of research work on ***C**ontext **M**odelling **f**or **A**daptive **A**uthentication systems* (CM4AA). To this end, we pursue three goals based on the *Systematic Mapping Study (SMS)* and *Systematic Literature Review (SLR)* research methodologies. We first present a SMS to structure the research area of CM4AA (**goal 1**). We complement the SMS with a SLR to gather and synthesise evidence about context information and its modelling for adaptive authentication systems (**goal 2**). From the knowledge gained from goal 2, we determine the desired properties of the context information model and its use for adaptive authentication systems (**goal 3**). Motivated to find out how to model context information for adaptive authentication, we provide a structured survey of the literature to date on CM4AA and a classification of existing proposals according to several analysis metrics. We demonstrate the ability of capturing a common set of contextual features that are relevant for adaptive authentication systems independent from the application domain. We emphasise that despite the possibility of a unified framework, no standard for CM4AA exists.

CCS Concepts: • **Context-aware systems** → **Adaptive Authentication Systems**; • **Context Modelling** → *Authentication Context*; • **Usability, Security** → *Usable Security*.

Additional Key Words and Phrases: Adaptive authentication, context information, user behaviour, systematic literature review, systematic mapping study

ACM Reference Format:
Anne Bumiller, Stéphanie Challita, Benoit Combemale, Olivier Barais, Nicolas Aillery, and Gael Le Lan. . On Understanding Context Modelling for Adaptive Authentication Systems. , ( ), 36 pages.

Authors' addresses: Anne Bumiller, Orange Labs, University of Rennes 1 / IRISA / INRIA, Rennes, France, anne.bumiller@orange.com; Stéphanie Challita, University of Rennes 1 / IRISA / INRIA, Rennes, France, stephanie.chalitta@inria.fr; Benoit Combemale, University of Rennes 1 / IRISA / INRIA, Rennes, France, benoit.combale@inria.fr; Olivier Barais, University of Rennes 1 / IRISA / INRIA, Rennes, France, olivier.barais@inria.fr; Nicolas Aillery, Orange Labs, Rennes, France, nicolas.aillery@orange.com; Gael Le Lan, Orange Labs, Rennes, France, gael.lelan@orange.com.

1 INTRODUCTION

In computer security, we mainly consider two forms of authentication: authentication of entities (a human user or a computer is who she claims to be) and authentication of messages (a message originates from the claimed sender) [13]. This work focuses on human entity authentication. We define an **entity** as a human that has a distinct existence, and that can be identified in context. We define **authentication** as *"the process of proving that an entity is genuinely who this entity claims to be "* [21]. This is a commonly used definition in the research field [14, 32].

Authentication is the ability to prove that an entity is genuinely who this entity claims to be and not necessarily a question of proving a unique identity (**identification** [21]). For example, a company service may only be accessible to employees. This means that the entity claims to be an employee. Authentication here comprises the process of verifying that the entity is an employee, whereas identification means to verify the unique identity of the employee.

Besides identification, **authorisation** also needs to be delimited from authentication. Authorisation is the process of verifying what specific resources an entity has access to [21]. Hence, in the example, it means to verify what the employee has access to. Authentication is about the question of who the entity is and authorisation about the question of what permissions the entity has. Literature on authorisation is not covered in this study. Authorisation is orthogonal to authentication and normally takes place after it [5]. Therefore, existing authorisation approaches can be integrated with adaptive authentication systems.

Authentication mechanisms require entities to provide claim information when they try to access resources in an information system or other **authentication targets**, such as services, devices, or systems. An **authentication system** is a system that uses authentication mechanisms in order to prove that an entity is genuinely who this entity claims to be.

Finding the balance between desired properties of such systems (*e.g.,* usability, security) is challenging. For this aim, the context needs to be taken into account so that the authentication mechanism can be chosen accordingly. For example, the geolocation of an entity may influence the need to verify the legitimacy of the entity. A deviation from habits, such as an authentication attempt from another country, can be due to the fact that the authentication attempt comes from an intruder situated in another country than the legitimate entity. Assuming an entity is situated at his workplace according to his habits, then an authentication challenge could be unnecessary and only disrupts the process. The role of **adaptive authentication** is to balance desired properties of the authentication mechanism (*e.g.,* security, usability) [6].

Let us consider the following example to illustrate the role of adaptive authentication. Bob, a German traveller in Spain checks his e-mails at 2:00 am in a poorly lit room. He enters the username and password correctly. His e-mail provider can acquire contextual information: geolocation, luminosity, time, and typing speed. Bob's e-mail provider determines some threats: Bob is not located in Germany as usual, he is checking his e-mails at an unusual time, it is dark around him, and he is typing slower than usual. All these threats make the e-mail provider assume that there is a risk that an intruder who has Bob's password might try to access Bob's e-mails. Bob has registered facial recognition and fingerprint as authentication mechanisms. Password-based authentication can be bypassed by the intruder who has stolen Bob's password. Face recognition is not efficient to use in the dark. Therefore, the adaptive authentication mechanism used by the e-mail provider determines that Bob needs to be authenticated with his fingerprint.

To enable authentication systems to take advantage of the context, a clear understanding of what **context** means is necessary. Dey et al. [12] propose a definition, which is also taken up by other authors working in the field of context modelling [4, 29, 57]: *"Context is any information that can be used to characterise the situation of an entity."*

Within this article, we shed light on the entities and their situations in an adaptive authentication system. A **context-aware system** is defined by Dey et al. [12] as *"a system that uses context to provide relevant information and/or services to the user, where relevancy depends on the user's task"*. According to this definition, we define an **adaptive authentication system** as a **context-aware authentication system** that uses context to provide the relevant authentication mechanism(s), where relevancy depends on the desired properties of the authentication mechanism for a user in a context.

Our work is related to Arias-Carbacos et al.'s survey on adaptive authentication [5]. In [5], the authors outline how to apply the design principles known in adaptive systems to adaptive authentication systems but do not deeply study context modelling and how the context information model is used in the authentication system. Complementary to [5] and leveraging on their conclusions, in this work we focus on context modelling for adaptive authentication systems and do not discuss self-adaptive systems design in general. Until now, context modelling for security applications (*e.g.,* adaptive authentication) has not been deeply studied [22]. In [5], the authors mention that most of the works surveyed in their article "show a limited usage of context, with vague descriptions and grounds". Leveraging on this conclusion, we conduct efforts to find out what models are suitable for the field of context modelling for adaptive authentication. Our study is an important first step towards less vague descriptions and grounds of using context for authentication systems. Hence, our work is complementary with [5].

Commonly the term **continuous authentication** is defined as a means of proving the identity of an entity based on context information in a passive manner [5]. The terms adaptive and continuous authentication are not always clearly separated from each other. According to our definition of adaptive authentication systems, we focus on providing the relevant authentication mechanism(s) regarding context information. We do not differentiate between active and passive authentication mechanisms and hence do not differentiate between continuous and non-continuous authentication mechanisms in our study about CM4AA.

Developing context-aware authentication systems need to be supported by adequate context information modelling techniques to reduce their complexity and improve maintainability [9]. We aim to support adaptive authentication practitioners on CM4AA. Therefore, we follow the procedures of the *Systematic Mapping Study* (SMS) and *Systematic Literature Review* (SLR) methodologies [42]. We achieve **three complementary goals**. The former one (SMS) enables us to structure the research area and to get a comprehensive overview of the research topic of CM4AA (**goal 1**). The latter one (SLR) enables us to gather and synthesise evidence about context information, it's modelling for adaptive authentication systems, and the use of the context information model (**goal 2**). The knowledge gained from goal 2 enables us to determine the desired properties of the context information model and its use for adaptive authentication systems (**goal 3**). In addition, we provide an analysis of industrial needs in form of an expert survey and a list of commercial adaptive authentication solutions.

The rest of this paper is organised as follows. We present our research questions in Section 2. In Section 3, we present an expert survey and industrial solutions for adaptive authentication. Our review methodology is presented in Section 4. In Section 5, we present the metrics and findings related to RQ1, in Section 6 those related to RQ2 and in Section 7 those related to RQ3. In Section 8, we assess strengths, weaknesses, opportunities, and threats of the research field of CM4AA. Threats to the validity of our study are discussed in Section 9. We present related surveys in Section 10. We conclude our work in Section 11.

2 RESEARCH QUESTIONS

In our work, we aim to analyse how context information modelling for adaptive authentication systems is performed to support adaptive authentication practitioners on CM4AA. Therefore, we aim to identify relevant publications on CM4AA to characterise what is the nature of the current body of knowledge about CM4AA (goal 1). We shed light on which context information determines the context of adaptive authentication systems and how it is modelled (goal 2). Also, we figure out which are the desired properties of the context information model and its use for adaptive authentication systems (goal 3). The three goals manifest in the three following research questions:

- **RQ1:** What is the nature of the current body of knowledge about CM4AA?
  The main activities to answer are:
  (1) to uncover which keywords and concepts reflect the research area of CM4AA to understand the nature of the research area and the notations in the domain,
  (2) and gaining an overview of the distribution of works in the research field of CM4AA regarding the year of the publication, the application domain, and the type of the contribution to understand the structure of the research area, when, how and from which point of view the research is conducted,
- **RQ2:** Which context information determines the context of adaptive authentication systems, how is it modelled, and for which phase of the authentication system life-cycle is the model used?
  The main activities to answer are:
  (1) establishing a holistic overview of which context information determines the context of adaptive authentication systems,
  (2) analysing context modelling approaches for adaptive authentication systems in the literature to date to understand the data structure according to which the context information model is built,
  (3) and analysing the use of the context information in the authentication system life-cycle.
- **RQ3:** Which are the desired properties of the context information model and its use for adaptive authentication systems?
  The main activity to answer is:
  (1) to uncover the desired functional and non-functional properties of the context information model and its use for adaptive authentication systems.

Fig. 1 visualises the relation between our three research questions and how we use the methodologies SMS and SLR to solve them.

3 INDUSTRIAL NEEDS

We aim to support adaptive authentication practitioners on CM4AA. Therefore, we designed a survey to uncover experts' thoughts on adaptive authentication and analyse adaptive authentication approaches applied in the industry.

3.1 Expert Survey

Our survey questions concern the **context information that can be used for authentication** (1) and **desired properties of adaptive authentication systems** (2).
We ask the experts question about whether and how context is used for authentication and what are desired properties of an authentication system. Table 1 shows some of the questions for our two question types. The totality of questions and anonymous answers can be found on our companion website (https://annebumiller.wixsite.com/slrcontext).

Table 1. Survey Questions

Context information that can be used for authentication	Desired properties of adaptive authentication systems
Is contextual information used to decide the authentication path in current authentication systems that you are using ? What contextual information is used during the authentication process? How do you rate the relevance of the following contextual information for authentication: device, IP address, web browser, geolocation, luminosity, time, user habits, nearby people, user activities (1-10)? Do you think that contextual information is used sufficiently during the authentication process? Why is contextual information not used in the authentication process? Is contextual information used for purposes other than authentication? Do you think it would make sense to use this same contextual information during the authentication process?	How is the suitability of an authentication mechanism assessed in a user path? What properties of authentication mechanisms are used to evaluate authentication mechanisms (usability, security, deployability, privacy)? Is the authentication pathway designed to address identified risks? Why are no risks taken into account when designing the authentication pathway? Do you think it would be appropriate to assess the risks during the authentication process and modify the process? What risks should be taken into account when designing the authentication path? What authentication mechanisms are offered to the user? Do you think that sufficient authentication mechanisms are currently available?

The Expert Panel. The expert panel consists of eleven people working on identity management, authentication, and system security. They come from a multinational telecommunications corporation (Orange), a multinational aerospace corporation (Airbus), two European university research institutes (University of Hohenheim, Chouaïb Doukkali University El Jadida), and a medium-sized family-owned company for smart sensor and image processing technologies (Wenglor Sensoric). We targeted people aware of the opportunity to use context information for authentication. It is not possible to identify and survey this entire population. Hence, we have chosen people from our professional network. All those people are potential adaptive authentication system designers and, therefore, potential users of our framework. Table 2 shows the job titles of the experts.

Table 2. Experts Job Titles

	Job Title
Expert 1	Identity Transverse Architect
Expert 2	Architect for Access Platforms
Expert 3	PhD Student: Behavioural Biometrics
Expert 4	Project Manager: Adaptive Authentication
Expert 5	System Architect of the Digital Identity Train
Expert 6	Direction of the Identity and Trust Research Program
Expert 7	Architect for Projects for Identity Anticipation and Research
Expert 8	Head Of Identity and Access Management for Users
Expert 9	Professor (Chair of Information Systems)
Expert 10	Master student of Big Data Analytics and Biometrics
Expert 11	Team Leader IT-Infrastructure

The Survey Procedure. In the first stage, the main idea of using context information (defined as any information that can be used to characterise an authentication attempt) for authentication was presented to the expert panel, followed by instructions on answering our online survey using a web questionnaire tool. Online survey is a faster way of collecting data from the respondents as compared to other survey methods like interviews. In addition, we invited the experts to contact us in the case of any questions or if they are interested in having an in-depth discussion. In the second stage, the experts answered our two question types (context information that can be used for authentication (1) and desired properties of adaptive authentication systems (2)). Three of the experts contacted us to discuss the topic further.

Analysis of the Responses. We analysed the experts' responses to our survey questions together with the interviews with the three experts with whom we had a detailed discussion. Most of the experts claim that **context information is not sufficiently used for authentication**. Nine out of eleven experts agree that context information is used for authentication, but eight of them claim that it is not sufficiently used. The two experts claiming that context information is not used mention the reason that there is a "lack of knowledge about how to use it". Hence, experts need more support to use and model contextual information for authentication. Furthermore, the great diversity of answers to the question of which context information is used (*e.g.,* device, risk score, localisation, browser fingerprint) shows that needs and perceptions vary greatly. This also points to the need for our study on CM4AA.
Five of the experts claim that the authentication path is the same for every authentication path of a user. This points out that nearly 50% of the experts think that there is not enough adaptation. The six experts claiming that there is an adaptation think that the authentication path is adapted to the sensitivity of the accessed resource, the availability of authentication mechanisms for a user, the contextual risks or to contextual information in general. This shows that the experts do consider adaptation at different levels and that notions are not unified in the domain. Support for using and modelling contextual information to allow adaptation is necessary.
Finally, ten out of eleven experts claim that **not enough authentication mechanisms are used**. At least five experts consider each of the properties: security (9), deployability (5), usability (10), and privacy (9) essential for an adaptive authentication system.

Results. Our survey results show that the experts need support to take full advantage of context information for authentication. We show that the experts are interested in **using contextual information** and do not yet make sufficient use of it. The **adaptation** of authentication decisions also interests the experts, and they find that this is not yet being done sufficiently. The properties **security**, **usability**, **deployability**, and **privacy** of adaptive authentication systems are considered important by the experts. Our study helps adaptive authentication practitioners to better understand context modelling for adaptive authentication systems.

3.2 Adaptive Authentication Applied in the Industry

In [56], the authors analyse risk-based authentication "applied in the wild" and determine the contextual feature set used during user login by LinkedIn, Facebook, Google, Amazon and GOG.com and derive how the adaptive authentication is applied in practice.
Furthermore, we searched for commercial adaptive authentication solutions. With the help of *Expert Insights* (https://expertinsights.com/), a cybersecurity research and review website, we identified common solutions. *Expert Insights*

provides guides, expert advice and industry insights to to help organizations to make informed, decisions when selecting cybersecurity solutions. They propose a list of top adaptive authentication solutions[1].

Prove Multi-Factor Authentication (MFA). Prove offers multi-factor authentication solutions that use users' mobile phones and phone numbers (phone-centric authentication) as the primary authentication method. The solution verifies a consumer's identity and validates the information provided by the consumer, assigning a trust score to each login to assess risks. The solution analyses behavioural and phone-related indicators of suspicious activity[2].

Duo. Duo offers MFA and *Single-Sign-On* (SSO) to allow access while only verifying once the identity. Administrators can configure adaptive authentication policies based on the user's location, device and role, among other factors. Duo then scans these security policies for anomalous access attempts to securely enable or deny access[3].

IBM Security Verify Access. This solution supports user authentication via one-time passwords, email verification and knowledge-based questions, and enables password-less SSO. Using the risk scoring engine, administrators can configure risk-based authentication policies to prevent anomalous login attempts. The risk scoring engine analyses the login patterns of users, including information about their devices and regular session activities to detect and prevent unusual login attempts[4].

Kount Control. Kount Control uses an AI-driven technology to analyze user login behavior based on device status, IP address reputation, geolocation and mobile and proxy indicators. Using this data, Kount detects anomalous access attempts that could be the result of attacks. In the case of a high-risk login, the system requires the users to verify their identity via an additional authentication method[5].

LastPass MFA.. LastPass MFA is an adaptive solution that combines contextual information such as geolocation and IP reputation, with biometric information, in order to analyze a user's risk score and verify their identity[6].

Okta Adaptive Multi-Factor Authentication. Okta Adaptive Multi-Factor Authentication uses contextual factors such as device trust and geolocation to calculate a risk score for login attempts before prompting users to further verify their identity. The platform supports secondary authentication via mobile app push notifications and biometrics, as well as more traditional methods, including security questions and *One-Time-Password* (OTP)s sent via SMS, phone call and email[7].

OneLogin SmartFactor Authentication. The solution aims to adjust authentication requirements in real-time based on the risk level associated with the context of each login attempt. The engine calculates risk scores based on user location, device security and user behavior, in order to determine the most appropriate action for each login to allow, deny or challenge the login by requesting up further verification. SmartFactor Authentication supports SMS, email and voice OTPs, security questions, push notifications via an app, and biometrics[8].

[1] https://expertinsights.com/insights/the-top-10-risk-based-authentication-rba-solutions/
[2] https://www.prove.com
[3] https://duo.com
[4] https://www.ibm.com/fr-fr
[5] https://kount.com/products/kount-control/
[6] https://www.lastpass.com/fr
[7] https://www.okta.com
[8] https://www.onelogin.com

Ping Identity PingOne Risk Management. The solution uses machine learning models to learn each user's login behavior, analysing risk predictors such as device type, operating system, browser version, date and time to distinguish between normal user login behavior and anomalous login attempts. Authentication policies that enable the system to grant, deny, or challenge access can be implemented based on a risk score calculated using the data[9].

SecureAuth Identity Platform. SecureAuth's Identity Platform utilizes artificial intelligence to produce a risk score for login attempts based on contextual information, such as device health, location, IP reputation and user behavior. If the risk associated with a login attempt is too high, SecureAuth will request further verification from the user[10].

Name	Self-designation	Context	Approach
Prove	MFA	behavioural, phone-related information	Trust score assignment to every authentication attempt
Duo	SSO	geolocation, device, role	Detection of anomalies based on contextual factors
IBM Verify Access	SSO	login patterns, session activities	Risk scoring engine to prevent anomalous logins
Kount Control	AI-Driven Solution	login behaviour, device, IP reputation, geolocation, mobile- and proxy indicators	AI-based anomaly detection
LastPass	MFA	geolocation, IP reputation, biometric information	Risk score calculation based on context
Okta	MFA	device, geolocation	Trust scores for device and geolocation
OneLogin	Access Management Solution	geolocation, device, behaviour	Risk score calculation based on context
Ping	Risk Management Solution	device, operating system, browser version, date, time	AI-based use behaviour analysis for anomaly detection
SecureAuth	AI-Driven Solution	device, geolocation, IP reputation, behaviour	AI-based risk score calculation

Table 3. Overview of Industrial Solutions for Adaptive Authentication

In summary we observe that industrial solutions are mainly aim to **assessing the risk or, conversely, the trust in the user** often based on AI and machine-learning technologies to calculate risk scores and to detect anomalies and derivations from user patterns. Table 3 summarises the different solutions. The providers call themselves by different names, although the approaches are all quite similar. There is a lack of standardisation.

In the analysis of industrial needs, we found that (1) there are commercial solutions for adaptive authentication, (2) that they are mainly based on the calculation of a risk score, and (3) experts need more support to model context. These results point out to the need of a study on CM4AA to support experts on context modelling and to allow more extensive approaches that only the consideration of a one-dimensional risk score.

[9] https://www.pingidentity.com/en.html
[10] https://www.secureauth.com

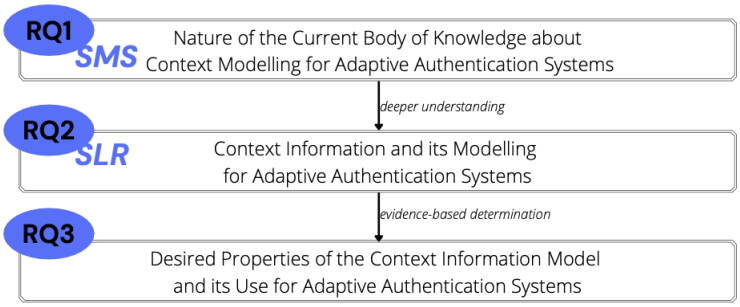

Fig. 1. Research questions and methodological approach to answer them

4 SYSTEMATIC REVIEW METHODOLOGY

In this section, we present our methodological approach based on the procedures of SLR and SMS [42][11] (Fig. 1). Within RQ1, we aim to structure the research area of CM4AA to understand the nature of the current body of knowledge about CM4AA. According to [42], SMSs are used to structure a research area, while SLRs are focused on gathering and synthesizing evidence. Hence, for solving RQ1, we apply the procedure of a SMS, and for solving RQ2, that of a SLR. Findings about the nature of the current body of knowledge about CM4AA (RQ1) allow us to understand and interpret those related to RQ2. With the help of the findings related to RQ2, we can determine the desired properties of the context information model and its use for adaptive authentication systems (RQ3).

In the following subsections, we describe our methodology to conduct the SMS and the SLR. We introduce the structure of our reusable search clause in subsection 4.1 and explain the exclusion criteria applied to the raw search results in subsection 4.2.

4.1 Logical Search Clause

We first analysed the recent literature in top academic venues and exchanged with domain experts (people working on identity management, authentication, and system security (see Section 3)). We used the snowball method to find literature by using the first references. Hence we obtained a set of representative papers to derive key terms.

Our search clause, consisting of a cartesian product of the terms presented in Table 4, is applied on GoogleScholar, ACM Digital Library, IEEE, Scopus, and SpringerLink. Essentially our search clause is a conjunction of the term "authentication system", "context modelling" and a disjunction of terms expressing the adaptation capability of the authentication system elicited after an initial scan of the literature published. For terms expressing the adaptation capability of authentication systems, we leveraged on the terms used in [5]. Thank to a snowballing approach, we assessed that "reinforced authentication" [17], "context-aware authentication" [19], "context-based authentication" [33], "progressive authentication" [47], "risk-based authentication" [56] and "risk-aware authentication" [20] are used in the literature appropriately to express the adaptation capability. Publications contributing to CM4AA need to use at least one of these terms. We included the spelling "context modeling" for "context modelling", the spelling "context-aware" for "context aware", the spelling "context-based" for "context based", the spelling "risk-aware" for "risk aware" and the

[11] All supplementary material (figures, tables with raw search results) is available on our cmpanion website: https://annebumiller.wixsite.com/slrcontext.

"authentication system"	"adaptation" "adaptive" "reinforced" "progressive" "risk-based" "risk based" "risk-aware" "context-based" "context based" "context-aware" "context aware"	"context modelling" "context modeling"

Table 4. Representation of our Logical Search Clause

spelling "risk-based" for "risk based". Authorisation is the process of verifying what specific resources an entity has access to. Hence, we do not include works focusing on "context-aware authorisation".

We restricted the scope to papers that contain "authentication system", because we only want to analyse modelling approaches where the context information is modelled for an authentication system and hence with the purpose of using the information for authentication. After an initial literature scan, we observed that papers that do not contain the term "authentication system" but only the term "authentication" often discuss authentication as a security aspect of a context-aware application, but the context is not modelled for the purpose of authentication (*e.g.,* [1]). In order to find out in which form context is represented so that it is suitable for authentication systems, we want to exclude such papers.

We searched for parts of the query separately (full text search) and joined the results manually to deal with the lack of support of complex clauses. We downloaded the citations in multiple parts and fused the results afterward.

Search Results. To mitigate sampling and publication bias, we conduct searches on formal databases (*e.g.,* ACM Digital Library) and indexes (*e.g.,* GoogleScholar). The raw search results of our logical search clause contain 111 publications:

- **GoogleScholar:** 69
- **IEEE:** 9
- **SpringerLink:** 16
- **Scopus:** 15
- **ACM Digital Library:** 2

We deleted 31 duplicates in the first step. We classified the remaining 80 publications according to the exclusion criteria described in the following section. Fig. 2 visualises our publication selection procedure. The publications of the type review, or study are helpful to gain background information on CM4AA and to analyse the year of publication and the contribution type, but the other analysis metrics have only been applied to contributions of the type concept, method, and tool (24 papers).

4.2 Exclusion Criteria

Based on common inclusion and exclusion criteria for systematic literature reviews proposed by the University of Melbourne[12], we determine the exclusion criteria for our work:

[12] https://unimelb.libguides.com/sysrev/inclusion-exclusion-criteria

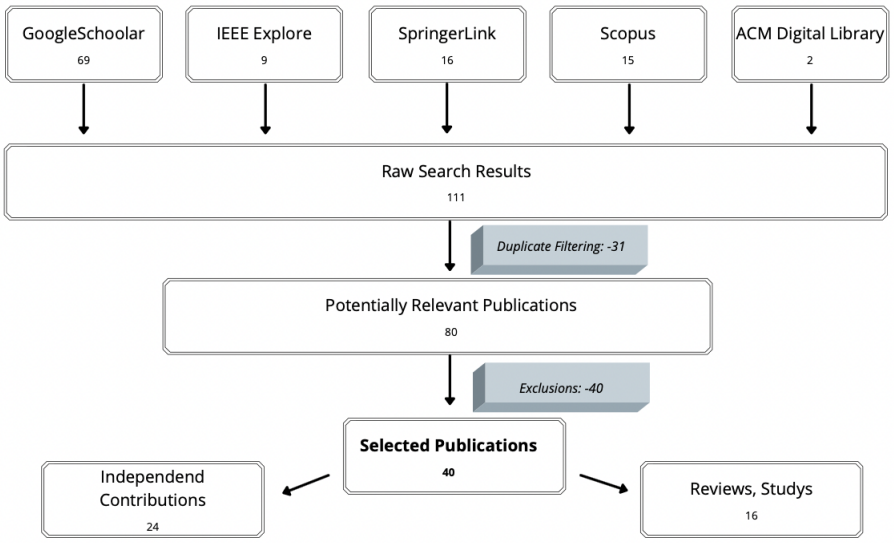

Fig. 2. Publication Selection Procedure

Table 5. Number of Publications per Year

Year	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Relevant Publications	4	2	1	3	3	3	8	2	5	5	4

- The paper is not in English.
- The paper is not accessible electronically.
- The paper is a short paper ($\leq$ 4 pages) or a teaser.
- The paper is a patent.[13]
- The journal/conference/workshop is not international.

Retaining papers per year. After having deleted the duplicates and having applied the exclusion criteria, we kept **40** publications for further analysis. Table 5 shows the number of kept publications per year from 2011 up to now. Fig. 3 shows the course of publications over the last 10 years and shows a continuous interest in the research area of CM4AA with a peak in 2017. Some fluctuation in the number of publications across different years can be observed but the interest in the topic always exist. The problem does not seem to be solved.

4.3 Analysis Process

For each research question (Section 2), we consider several metrics to analyse the publications. First, all six analysts worked together to determine which raw data is needed for each metric. Second, we have divided the papers among ourselves (six subsets) and each analyst collected the necessary raw data from a subset of the reviewed papers (manual extraction after reading). Third, we analysed the data according to the metric (*e.g.*, classification, frequency of occurrence). For this, each analyst has analysed a subset of papers. For a set of 10 papers, all the six analysts conducted the analysis

[13]Patents are excluded from further analysis, but the high number of existing patents shows industrial interest in the topic and suitability of the research domain for industry.

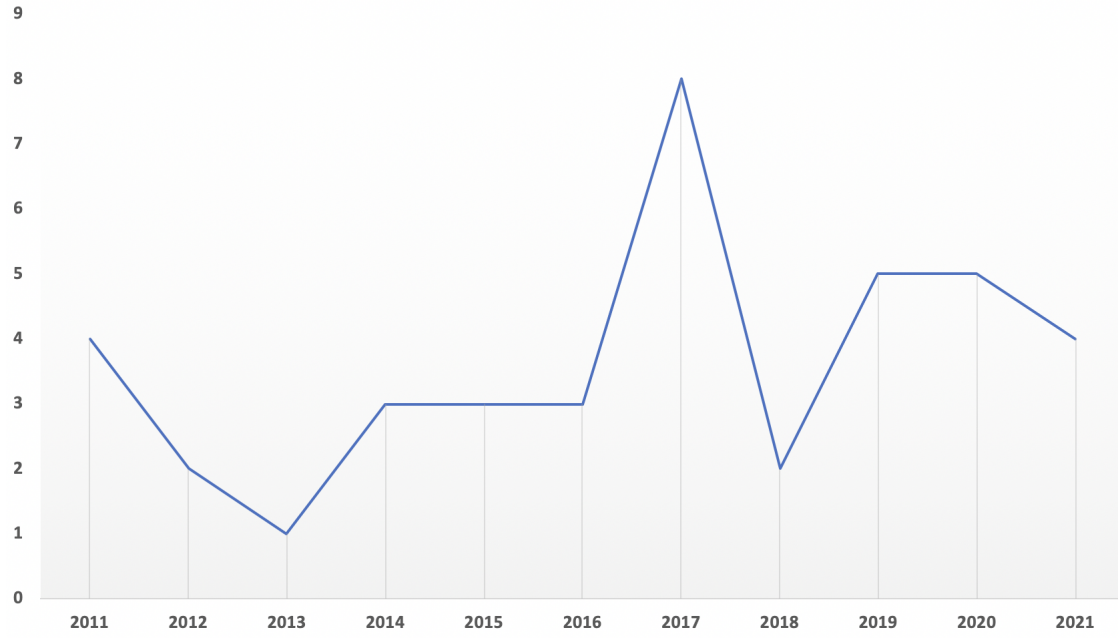

Fig. 3. Course of Publications Over the Last Ten Years

independently and discussed the results all together. This discussion served to align the typical answer types and share a common understanding regarding the different criteria. For the other papers, at least two experts did the analysis and discussed the results. Three of the analysts are experts in the field of adaptive authentication, the other three are experts in the modelling domain. In regular synchronisation meetings we discussed our analyses. We solved conflicts according to the majority principle if it was possible. If not, we asked another reviewer to read the paper and make a decision.

5 RQ1: NATURE OF THE CURRENT BODY OF KNOWLEDGE ABOUT CONTEXT MODELLING FOR ADAPTIVE AUTHENTICATION SYSTEMS

RQ1 concerns the nature of the current body of knowledge about CM4AA. In particular, we aim to better understand the research field of CM4AA, such as which keywords and concepts reflect the research field, what is the distribution of works concerning the year of publication, the application domain, and the type of contribution to better appreciate the nature of the findings in the following research questions.

5.1 Metrics for the Publication Analysis

We apply the methodology of a SMS to structure the research area of CM4AA. We present in this section the metrics considered to analyse the relevant publications.

5.1.1 Main Keywords. We aim to uncover which keywords and concepts reflect the research area of CM4AA.

Raw data. We collect the titles, the abstracts and the author-specified keywords (if available) for the selected papers.

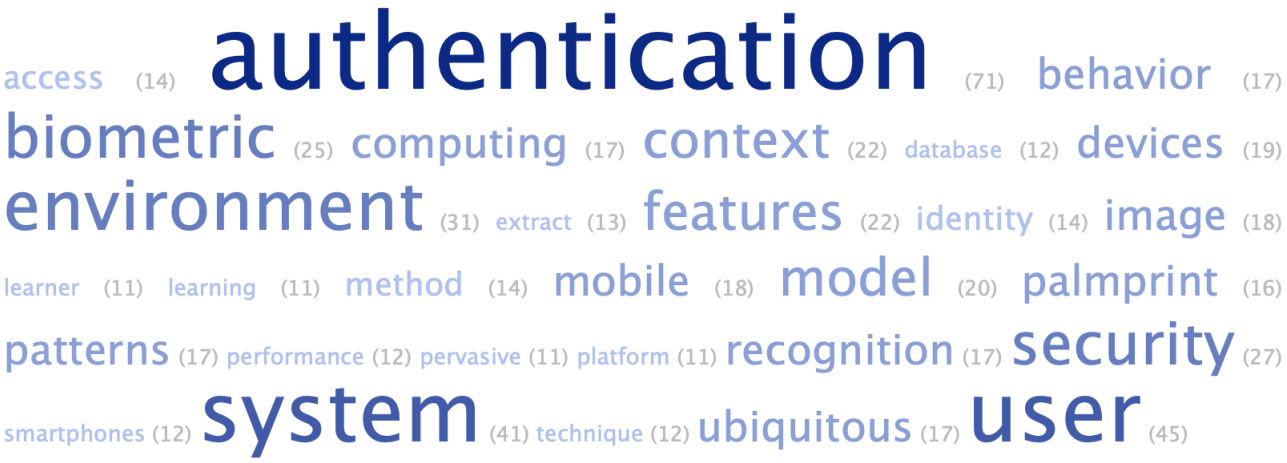

Fig. 4. Word Cloud Keywords - Titles, Abstracts, Author-Specified Keywords

Metric. Based on the raw data collected from each article, we filter the common keywords[14] and calculate the frequency of appearance of each word based on Stem algorithm [45]. The 30 keywords that appear the most often in the abstracts, titles and author-specified keywords of the publications are assumed to be the main keywords in the research field. The title and the abstract of a publication are usually the first introductions readers have to the work and therefore contain the main concepts. Additionally authors specify keywords that mostly reflect their work. We think that 30 is a reasonable number because with a larger number, the words are repeated (synonyms), and with a smaller number, only the ones from the search clause are repeated. The keywords are visualised in a word cloud (Fig. 4). As a visualization tool, we use TagCrowd[15], because of its ease to read, analyse and compare[16].

5.1.2 Contribution Types. We aim to uncover how research is conducted in the research area of CM4AA.

Raw data. We classify the publications along the type of research they conduct to understand how research is performed in the field of CM4AA. We classify the contributions based on [41] into concepts, methods, tools, studies, and reviews:

- **Concepts:** papers suggesting abstract ideas of how to model context for adaptive authentication systems by observing and analyzing already present information.
- **Methods:** development of concrete ways of CM4AA.
- **Tools:** papers presenting novel systems, prototypes, or software tools.
- **Reviews:** papers reviewing related literature.
- **Studies:** papers analysing and evaluating existing tools, methods or concepts.

One of the contribution types, concept, method, tool, review or study, is assigned to each of the reviewed publications. We did the assignment in a disjunctive manner: papers, suitable for more than one research type, were discussed and

[14] based on the following list https://tagcrowd.com/languages/English and according to our research goals
[15] https://tagcrowd.com/
[16] Additionally, we show the keywords in a table on our companion webpage.

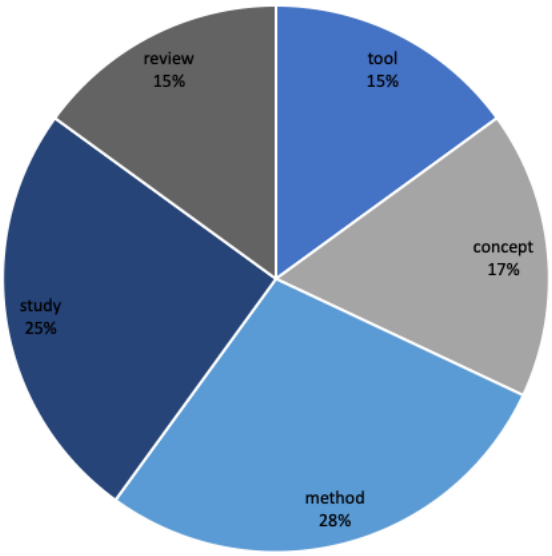

Fig. 5. Partition of the Contribution Types of the 40 Publications Relevant to this Article

assigned the most suitable contribution type. Here we consider the most suitable type to be the one at the focus of the contribution.

Metric. Fig. 5 is a pie chart that visualises the proportions of the contribution types.

5.1.3 Covered Application Domains. With the analysis of the application domains, which are covered in the field of CM4AA, we aim to uncover application domains in which CM4AA plays a crucial role.

Raw data. The application domain of a publication is the segment of reality (*e.g.,* telecommunication, healthcare, education) that is addressed within the publication. For each of the reviewed papers, we classify it according to its primary application domain if there is one or we indicate that the approach is generic.

Metric. After classifying the papers along the **years of publication** (Fig. 3), the **keywords** (Fig. 4 and the **contribution types** (Fig. 5), they are classified along the **application domain**, to enable the identification and discussion of domain-specific trends. An application domain is assumed to be covered if at least one contribution addresses the domain. 92% of the analysed publications are not specific to any application domain and can be applied to CM4AA in any domain. We identified two papers specifically relevant to the domain of education [18, 31].

5.2 Findings on the Nature of the Current Body of Knowledge about context modelling for adaptive authentication

We present in this section the findings on the nature of the current body of knowledge about context modelling for adaptive authentication.

5.2.1 Main Keywords. That the words *authentication* (71), *system* (41), *context* (22) and *model* (20) occur frequently is not surprising in regard of our search clause, but confirms the significance of our chosen terms. That *authentication*

appears more than twice as often as *model* can be interpreted as a clue that the research field of CM4AA is mainly authentication driven. The modelling community seems to have fewer contributions. This can also be seen as a reason for the lack of standardised context modelling methods for adaptive authentication systems. As we have explained, we focus on papers based on context modelling and explicitly exclude papers that deal only with authentication, and yet these seem to be driven by the authentication community.

Our search clause contains a disjunction of words expressing the adaptation capability of authentication systems. None of them is among the 30 most frequent words of the abstracts, titles and author-specified keywords of the papers. In a generic MAPE-K architecture for adaptive systems, there is one concern about gathering and representing managed resources and another concern about the actual adaptation logic. In an adaptive authentication system, the first concern refers to a capability to take into account the context information (context-awareness), while the adaptation logic refers to the capability of a system to change its behavior in response to the context. In this study, we target papers that focus on context-awareness and we observe that such works deal little or not at all with the actual adaptation logic.

The keywords *biometrics* (25) (*palmprint* (16)), *behaviour* (17) and *patterns* (17) show the trend of using these *features* (20) for adaptive authentication [34, 35]. *Databases* (12) from which the information can be *extracted* (13) seem to be important. The state of environmental elements (*environment* (31)) plays a role for adaptive authentication. Authentication is the ability to prove that an entity is genuinely who this entity claims to be (see Section 1) and not necessarily a question of proving an unique identity. When contextual features are used that confirm an unique identity then often the term *recognition* (17) is used. It seems to be common to use contextual features that clearly determine a unique **identity** (14). This justifies also the frequent appearance of the word *image* (18). In approaches working with images, those are often used to recognise biometrics (*e.g.,* palmprint, iris). In the works, the *performance* (12) of the approaches is often evaluated. *Platforms* (11) seem to be a relevant authentication target. The word *user* (45) indicates that the entity being authenticated is often the user. The frequent appearance of the word *security* (27) can be justified by the fact that authentication is an essential security aspect of systems [28]. *Smartphones* (12) and *ubiquitous* (17) *computing* (17) environments are important concepts in the research field of CM4AA. Context information acquirement with *mobile* (18) *devices* (19) is often easier than with non-mobile devices. Overall, biometric and behavioral information can be acquired more easily from mobile than from non-mobile devices. Anyway, non-mobile devices do not need to be neglected. The keyword *learning* (11) can be interpreted as a clue that the works often propose machine learning algorithms for adaptive authentication. The keyword *learner* (11) points out that education is a relevant application domain in the research area of CM4AA. *Access* (14) control is frequently used semantically similar to authentication. The terms authentication and access control are not always clearly separated from each other. We observe that terms that are clearly defined in the security domain (see Section 1) are not always used properly in the domain of CM4AA.

5.2.2 Contribution Types. There is a large number of studies and reviews (40%). Gaining an understanding of the existing research relevant to CM4AA seems to be in the interest of many researchers. The works fall in the categories of context and context-awareness, authentication modalities, adaptive authentication in specific computing environments and adaptive authentication in general. There is no review of works on context modelling for adaptive authentication systems. 15% of the contributions are of the contribution type tool. Adaptive authentication is a new research area and not yet every proposed concept of how to model context information for adaptive authentication systems goes beyond conceptualization and results in a tool. There are contributions of the type method (28%) and concept (17%). These works do not (yet) result in tools. CM4AA seems to be a conceptual and methodological research field. This research type, generally related to abstract ideas or schemes is a potentially powerful way to introduce new ideas,

to identify problems and appropriate solutions in new ways, and to provide new frameworks. Difficulties related to methods and concepts are the conflicts that may arise within the different approaches and their unsuitability for real world applications. Due to privacy and confidentiality issues, there is a lack of public authentication data, that would allow to push further the development of tools. For adaptive authentication system designers it is challenging to use context information efficiently without the support of tools.

5.2.3 Covered Application Domains. Most of the publications are not specific to any application domain (92%). This sheds light on the fact that CM4AA is a cross domain research topic. The danger is that terms are confused or concepts are understood differently. The right balance between desired properties of authentication mechanisms which is crucial in the context of adaptive authentication needs to be adjusted according to the domain. Based on the publications identified to be specific to an application domain, CM4AA seems to be particularly relevant in the domain of education. For online learning platforms it is crucial to adapt contents to the entities roles and needs. For example, students need, unlike teachers, not to have access to exam results. Anyway, it is possible that researchers who study CM4AA are teachers and therefore use the education application domain. However, this does not necessarily mean that education is a field of application in which CM4AA is particularly important.

Lessons Learned.
We observe a **continuous interest** in the research field of CM4AA over the last ten years. Works related to CM4AA focus on **context-awareness** and the actual **adaptation** capability of authentication systems is often disregarded. The research field is mainly driven from the **authentication** community. There is a trend of using **biometric** and **behavioural** contextual features that can be used to clearly identify a unique entity. It seems to be disregarded that authentication is not necessarily about proving a **unique identity**. In the research area of CM4AA, terms are not always clearly **delimited** from each other (*e.g.,* access control and authentication), what sheds light on the **lack of a standard** for CM4AA. **Mobile computing environments** and authentication on **mobile devices** are crucial in the research area of CM4AA. CM4AA is a **cross-cutting concern in multiple domains**, that integrates information from multiple disciplines or bodies of specialised knowledge. There are **concepts** and **methods** proposed in the literature that do not go beyond conceptualization and do hence not result in concrete **tools**. Due to **privacy** issues, there is a lack of public available data to push further the development of tools and benchmark solutions.

6 RQ2: CONTEXT INFORMATION AND ITS MODELLING FOR ADAPTIVE AUTHENTICATION SYSTEMS

RQ2 concerns context information and its modelling for adaptive authentication systems.

6.1 Metrics for the Publication Analysis

We gather and synthesise evidence about context information, its modelling for adaptive authentication systems, and the use of the model in the authentication system life-cycle within the methodology of a SLR and with the help of several analysis metrics.

6.1.1 Context Information. With the analysis of the context information that determines the context for adaptive authentication systems, we aim to uncover the context information which is most commonly used. We assume the context information to show up in a triplet *[Informing Entity, Contextual Feature, Assigned Entity]*, that allows us to analyse the entities and their situations in an adaptive authentication system in a detailed manner to be able to refer to

the definition of context information from Dey et al. [12] (*"Context is any information that can be used to characterise the situation of an entity."*). For example, the contextual feature location can originate from a smartphone and be attributed to a user: *[smartphone, location, user]*.

- **Informing Entities (IE).** Informing entities, such as devices or users, are entities that inform about the context. For example, a mobile device can inform about the contextual feature location.
- **Contextual Features (CF).** A contextual feature is a feature which is characterising the context of an entity (*e.g.,* its location, its behaviour). We consider contextual features coming up at two different **levels of transformation**. At the low transformation level (*e.g.,* raw sensor information like the location), and at the high transformation level (*e.g.,* information transformed from sensor information like an entity's behaviour).
- **Assigned Entities (AE).** Entities whose context is determined with the contextual features are entities the context is assigned to (*e.g.,* user, device).

Raw data. For each of the reviewed papers, we collect the information regarding the concepts of *IE*, *CF*, *AE* that appear within the publications. This information is directly extracted from the papers. We do not establish an a priori list of elements that can appear in this list. If an article does not discuss an element of this triplet, it is not classified in the corresponding category.

Metric. The metric for the three categories is a partition for each category of the frequency of occurrence of the collected items.

- Fig. 7 shows the partition of the most frequently **informing entities**. The device as IE means that the information is taken from the device (*e.g.,* integrated sensors). In some cases the information is directly taken from the environment (*e.g.,* with the help of a thermometer, light sensor). The system is assumed to be the IE when the system provides information directly (*e.g.,* diagnostic and troubleshooting information related to the operating system, hardware and software). Especially in the context of signal processes, images are used as input data to extract information. In some work, the user is assumed to inform about the context.
- Fig. 6 shows the partition of the most frequently used **contextual features**. Behaviour describes how an entity acts or conducts oneself (*e.g.,* typing behaviour), biometric describes biological measurements or physical characteristics (*e.g.,* fingerprint), activity describes the way in which an entity conducts towards the system (*e.g.,* requested resources), device information describes the piece of equipment which is used by the entity (*e.g.,* name of a mobile phone), environmental factors describe factors external to a person (*e.g.,* luminosity, background noise), location describes a particular place or position (*e.g.,* France), personal user information is any information related to an identifiable user (*e.g.,* address, phone number), roles describe an entities privileges (*e.g.,* administrator) and time the measured or measurable period during which the authentication attempt happens (*e.g.,* October, 10th 2021 at 09:09:09). We also calculate the percentage of papers which consider contextual information on a transformed level (*e.g.,* the behaviour) and not only on the raw sensor level (*e.g.,* the temperature).
- In 92% the user is the **assigned entity**. In the remaining works the context information is assigned to the device or the system.

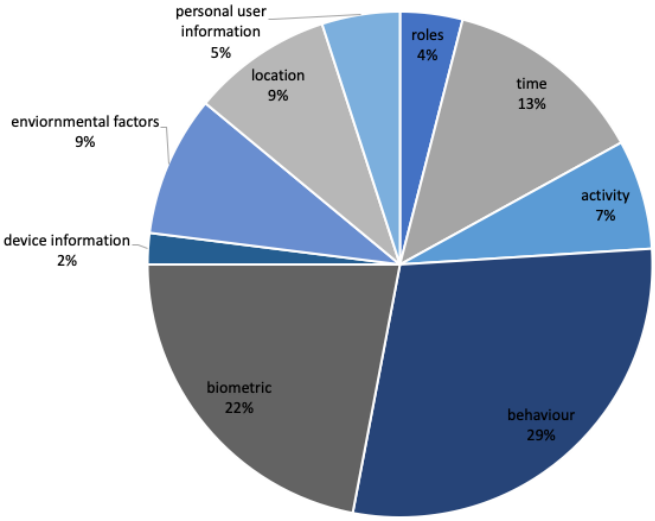

Fig. 6. Partition of the Most Frequently Used Contextual Features

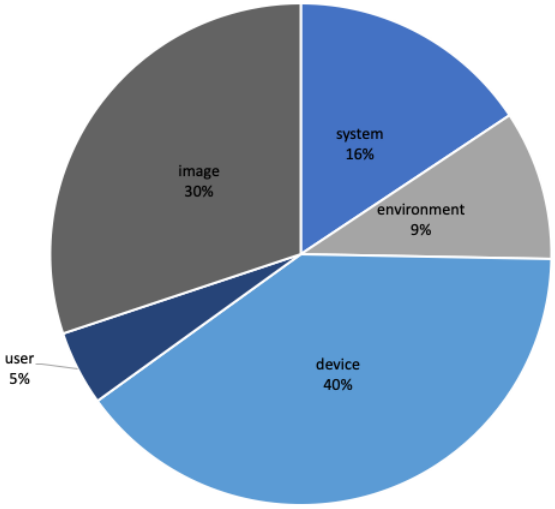

Fig. 7. Partition of the Most Frequently Used Informing Entities

6.1.2 Modelling Formalisms. We analyse the **modelling formalisms** for modelling the context for adaptive authentication systems proposed in the publications relevant to this article. We aim to uncover how context information modelling for adaptive authentication systems is performed to analyse how context models that are suitable for the field are defined and evaluated.

The modelling formalism consists of two parts:

(1) **Modelling Concepts.** The abstraction of the ideas and the definition of their precise meaning and relationships

(2) **Modelling Technique.** The technical approach (technological stack) according to which the model is built (*e.g.,* a standard modelling language). It defines the textual or graphical syntax of the model.

Raw data. For each of the reviewed articles selected, we analyse whether the introduced **modelling concepts** are generic, specific to an application domain or authentication specific:

- **Generic Concepts.** The concepts are generic if they are kept abstract and general, without ideas related to the authentication problem or a specific application domain (*e.g.,* contextual feature).
- **Authentication-specific Concepts.** The concepts are authentication-specific if they are related to the authentication problem (*e.g.,* authentication attack).
- **Domain-specific concepts** The concepts are domain-specific if they are related to a specific application domain (*e.g.,* learner for the education domain).

We identify the following four objectives on the basis of which the **modelling technique** is chosen:

(1) Mathematically formalize complex relationships
(2) Capture authentication security rules and threats
(3) Visualize the organisation and relationships among different functionalities of the system
(4) Represent processes in the authentication system

For each of the papers selected, we analyse the **modelling concepts** and the **modelling techniques**, we classify the **modelling concepts** into generic, authentication-specific and domain-specific concepts and the **modelling techniques** according to the underlying objective.

Metric. Fig. 8 shows the proportion of domain-specific (8%), authentication-specific (17%) and generic (75%) concepts that are proposed in the publications relevant to this article. The assignment is done in a disjunctive manner[17] depending on the starting point the authors propose for the modelling concepts: general concepts, domain-specific concepts, or authentication-specific concepts.
Fig. 9 shows the proportion of the underlying objectives of the used modelling techniques (Formalize mathematically complex relationships: 54%, Visualize the organisation and relationships among different functionalities of the system: 21%, Represent processes in the authentication system: 17%, Capture authentication security rules and threats: 8%).

6.1.3 Authentication System Life-cycle Stage. With an analysis of the distribution of the publications concerning the **authentication system life-cycle stage the context model is used for**, we aim to uncover lacks in existing context modelling approaches for adaptive authentication systems.

Raw data. The context model defines how context data are structured and maintained to produce a description of the context information that is present in the context-aware authentication system. There are three life-cycle stages of the authentication system: design (1), which is the phase of making design decisions regarding the architecture and structure based on gathered requirements and criteria, deployment (2), which is the phase of deploying the system in a production environment (configuring infrastructure, defining deployment strategy) and runtime (3), which is a representation of the authentication system that can be manipulated at runtime (the context information can be used at runtime) [8]. To structure and maintain the context information over the whole life-cycle of the authentication systems, concerns belonging to each stage should be considered in the model. We check for each context model identified in

[17] Papers that contain concepts from more than one category are assigned to the category that predominates.

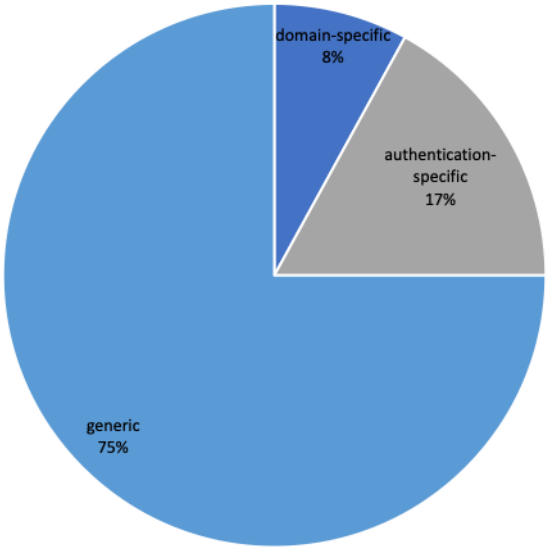

Fig. 8. Partition of Generic, Authentication-Specific and Domain-Specific Modelling Concepts

the literature for which stages it is intended and we we classify the models to belong to one or more system life-cycle stages.

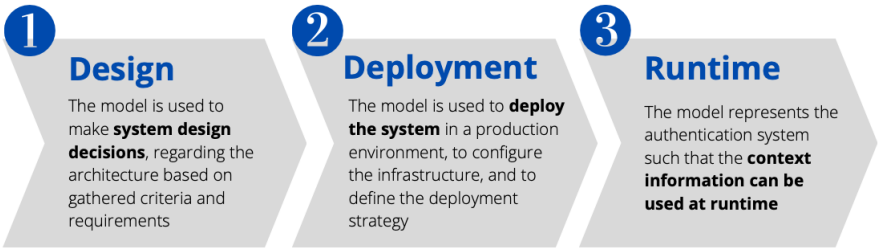

Fig. 10. Authentication System Life-Cycle Stages That the Context Model is Used For

Metric. Fig. 11 represents the proportions of publications relevant to this article that address the design-, the deployment- and the runtime-stage.

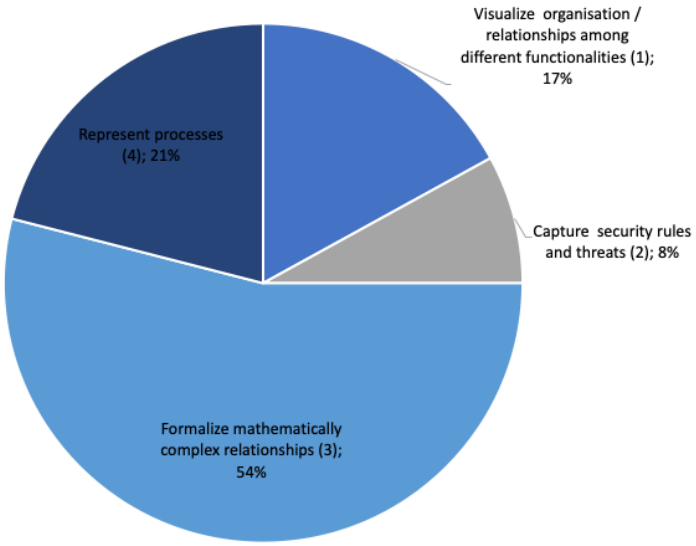

Fig. 9. Proportion of Underlying Objectives of the Proposed Modelling Techniques

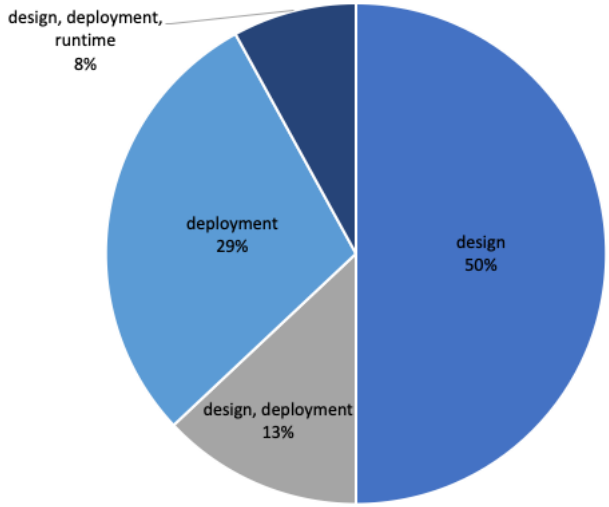

Fig. 11. Partition of the context models used for the design, deployment and runtime life-cycle stage of the authentication system

6.2 Findings related to Context Information and its Modelling for Adaptive Authentication Systems

In this subsection, we answer RQ2, we discuss which context information determines the context for adaptive authentication systems, how it is modelled, and how the model is used for adaptive authentication systems. The findings related to RQ1 show that CM4AA is a cross-cutting concern in multiple domains. Hence, we do not analyse domain-specific trends in this section, and we take into account issues related to interdisciplinarity. According to the findings related to RQ1,

biometric and behavioural information is commonly used for adaptive authentication in mobile computing environments. Hence, in this section, we treat issues related to these contextual features and mobile computing environments.

6.2.1 Context Information. Conform to the context information triplet, we analyse the informing entities, the contextual features, and the assigned entities in the following.

Informing Entities. We analyse which entities are informing about context information, and we discuss the data types and formats of the given context information. In 40% of the works, authors propose the use of context information which is acquired from sensors of mobile devices [36]. Mobile devices are crucial for data acquisition in the research area of CM4AA. The constant use of mobile devices has become a normality in our society. Hence, following this trend, authentication is increasingly discussed for mobile devices. This shift is also related to data acquisition: mobile devices are increasingly equipped with sensors, which makes the use of context information for authentication possible. This is an advantage, but it also brings new challenges to light, including the use of multiple devices in smart home and mobile computing environments. Despite the increased dominance of mobile devices, non-mobile devices must not be disdained either. Accelerometer, *Global Positioning System* (GPS) and touchscreen sensors are frequently used. Witte et al. [57] propose to automatically acquire the geolocation with the GPS sensor of a mobile device.
Images (30%) are crucial as well to inform about the context (*e.g.,* for the comparison of palm print images [26]). In 9% of the works the environment is informing about the context (*e.g.,* [29].
Depending on how the context information is used in the proposals, the data is represented in several data formats. Server logs [34] and time series [38] are popular formats, especially in works that are reasoning patterns and trends from the context information. In several works, the authors specify the data storage and discuss related issues. Often the data is stored in databases [29], in central repositories [40] or local repositories [50].

Contextual Features. Fig. 6 shows that the behaviour (29%) and biometrics (22%) are the most frequently used contextual features. In some works, the location is modelled for adaptive authentication systems (9%). Environmental factors, like nearby people or devices, the luminosity, or the noise, are often referred to as well when the context for adaptive authentication systems is modelled (9%). In their adaptive authentication system design methodology, Arias-Cabarcos et al. [4] propose taking into account the geolocation as a contextual feature. In the work from Ramakrishnan et al. [46] activities are modelled to detect anomalies. Neverova et al. [38] propose a method for active biometric authentication based on motion patterns.
61% of the contributions do not only rely on raw sensor data information (*e.g.,* location, temperature) but consider context information on a transformed level like the user's activities or behaviour.

Assigned Entities. In 92% of the reviewed works the **user** is the entity the context is assigned to (*e.g.,* [40, 48]). Ma et al. [31] assign the context information to **resources**. In other reviewed papers [27], the context information is assigned to the device. In the paper specific to the domain of education [18] the context information is assigned to the learner (domain-specific user).

6.2.2 Modelling Formalisms. We analyse the modelling concepts and the modelling techniques to understand how the context information is built for an adaptive authentication system.

Modelling Concepts. Most of the reviewed papers are not specific to any application domain and hence only 8% of the papers introduce domain-specific modelling concepts. In two papers education domain-specific modelling concepts are introduced [18, 31]. The fact that those papers that belong to a specific application domain (education) introduce

domain-specific concepts shows that formalising the authentication system structure, behavior, and requirements within particular domains is important.

The largest part of the identified modelling concepts are generic (75%). In this way, concepts are related to abstract types but do not require specific descriptions or relationships related to an application domain or the authentication problem. The fact that mainly generic concepts are introduced demonstrates the ability of capturing a common set of concepts and relationships for CM4AA. It is interesting to note that despite this possibility, no general standard for CM4AA exists.

There are also some authentication-specific modelling concepts (17%), which shows that CM4AA is driven by the authentication community.

Modelling Technique. We cannot identify a trend in the use of a particular syntax for CM4AA. Different structures to represent complex concepts and relationships visually or textually are presented in the reviewed works.

Nevertheless, four main objectives emerge: visualize the organisation and relationships among different functionalities of the authentication system (1), capture authentication security rules and threats (2), mathematically formalize complex relationships (3), and represent processes in the authentication system (4).

- (1) Visualize the organisation and relationships among different functionalities of the authentication system
  - **Component-based modelling**, which focuses on the decomposition of the model into individual components. It provides a higher level of abstraction and divides the problem into sub-problems (*e.g.,* context gathering and context analysis) [3, 18, 30, 57].
  - **Blockchain modelling,** which is a modelling approach based on an interlinked systematic chain of blocks that contains the history of data (*e.g., to take into account the history of contextual information*) [31].
- (2) Capture authentication security rules and threats
  - **Attack-Tree Modelling**, which deals with how vulnerabilities are exploited (*e.g.,* distinguishing between different attack types) [36].
  - **Rule-based modelling,** which is a modelling approach that uses a set of rules that indirectly specifies a model (*e.g.,* security rules) [52].
- (3) Mathematically formalize complex relationships
  - **Mathematical modelling,** which is a description of a system using mathematical concepts and languages (*e.g.,* the representation of context information in a vector) [4, 10, 15, 16, 25, 29, 34, 37, 38, 44, 46, 49].
  - **Biological modelling,** which is a modelling approach inspired by biological phenomena (*e.g.,* modelling context information as a Chromosome where each individual context is a gene) [50].
- (4) Represent processes in the authentication system
  - **Flowchart modelling,** which is a type of diagram that represents a workflow or process (*e.g., to model the reasoning about context information for adaptive authentication within a flow of steps*) [26, 27, 40, 48].

We see in Fig. 9 that many works (54%) focus on formalising mathematically complex relationships. Authors aim to exactly represent the real problem situations. We have already noted that approaches are often presented that clearly identify a single entity. This requires precise calculations and comparisons. (*e.g.,* for the comparison of palm print images [26]). For this purpose, a mathematical modelling syntax is well suited.

In 21% of the works, different functionalities of the authentication system are separated and represented in different model components. The models describe the components used to make the desired functionalities of the authentication system. Component diagrams can also be used to construct executables by using forward and reverse engineering.

In 17% of the reviewed works, system processes are described in the proposed model. Flowchart is is an important tool for planning and designing a new system, it provides an overview of the system and also demonstrates the relationship between various steps.
In 8% of the proposed modelling approaches the main objective is to capture security rules and threats. As authentication is an important security aspect of the system it is important to take into account such threats and rules.

6.2.3 Authentication System Life-cycle Stage. Within an analysis of the contributions regarding the life-cycle stage of the authentication system that the context model is used for, we aim to detect trends and gaps in the literature.
More than half of the publications (63%) focus on the **design** of the system. In these works, the context model serves as a representation that can aid in defining and analyzing a set of concepts of the adaptive authentication system. In [18] for example, the model serves as a representation of the concepts of learning system architecture without considering concerns about deployment or runtime. The concepts (*e.g.,* "service credential request") are used to analyse the authentication procedure. An overview of different functional components of the system are represented in the model in [46].
In 13% the design stage is addressed together with the deployment stage.
In 29% the **deployment**-stage is addressed. In those works the model is implemented but not used at runtime. In [29], the model representing the system architecture has additional modules that allow the system implementation.
In 8% of the works design, deployment and **runtime** issues are addressed. In these works the authors explicitly address the system execution. A common purpose for models at runtime is self-adaptation [8]. This is the case also for the works we identified that treat CM4AA at runtime. The fact that only a few papers deal with adaptation shows again that this aspect is not a major issue in the papers that deal with context modelling even if the ultimate end goal of an adaptive authentication system is necessarily to adapt at runtime.
We mentioned in Table 3 that existing run-time solutions are mainly based on the calculation of a one-dimensional risk score. Using the context information model at runtime for adaptive authentication systems in a more extensive manner is rarely studied.

Lessons Learned. Often the works are based on context information acquired from **mobile devices**. Those are therefore crucial for data acquisition in the research area of CM4AA. **Non-mobile devices** are often disregarded. The commonly used **context information** (biometrics, behaviour, location) is highly **privacy** sensitive information. This makes it difficult to ensure the user's willingness to disclose private context information even if it is used for the purpose of authentication. It is common to determine **patterns** and habits from the authentication history of users. This can be an advantage regarding the **storage** of the context information. In some cases, only the habits, like the usual location, need to be stored and not the whole history of authentication attempts. Regarding the **privacy** this can be an advantage as well. Other **anomalies** than derivations from patterns and habits are often disregarded. In works that focus on human identity authentication, the context is usually assigned to the entity which needs to be authenticated. That there are only a few works also considering contextual features assigned to other entities sheds light on the fact that the **contextual relations** between different entities often are omitted when context information for adaptive authentication systems is modelled. The largest part of the identified modelling concepts are generic (75%). We cannot observe a trend in the use of a **modelling technique** to model context information for adaptive authentication systems despite the clear identification of the underlying goals. There is a great diversity of syntax proposed in the literature, which sheds light on the lack of a modelling **standard** for CM4AA systems. This is also related to the fact that the research area of CM4AA is mainly authentication driven and the influence of the modelling community is limited. The lack of standards makes it difficult for adaptive authentication practitioners to model context information efficiently and structured. Also, standards would help to clarify **reglementations** regarding privacy issues, and users would be more willing to share context information if it is modelled according to an accepted standard and used for adaptive authentication in a regulated manner. The *National Institute of Standards and Technology* (NIST) proposes **guidelines** for authentication and the management of digital identities, which need to be used also in order to establish appropriate modelling standards. The context information models are mostly used at the **design time** (63%) and **deployment time** (42%) of adaptive authentication systems. There is a lack of works treating CM4AA systems at runtime (8%). The lack of works treating CM4AA at **runtime** is due to the lack of concrete implementations. Even if the end goal of an adaptive authentication system is to adapt at runtime, many research proposing context models for adaptive authentication systems actually does not address runtime concerns. Often there is no data available. Adaptive authentication is still a **young research area** and is not yet much applied at runtime. Runtime is when the application is running and not yet much complete adaptive authentication applications are running.

7 RQ3: DESIRED PROPERTIES OF THE CONTEXT INFORMATION MODEL AND ITS USE FOR ADAPTIVE AUTHENTICATION SYSTEMS

RQ3 concerns desired properties of the context information model and its use for adaptive authentication systems.

7.1 Metrics for the Publication Analysis

We do not identify a standard from which we can derive desired properties on the context information model and its use for adaptive authentication systems. Nevertheless, the authors of the reviewed papers identify constraints on how context information modelling is done successfully for adaptive authentication systems. We observe that various properties has been identified as important for the context model to be suitable for adaptive authentication systems. Some of these constraints are also evaluated empirically in the reviewed works. In order to understand which properties the authors consider important, we perform an analysis of these constraints.

Raw data. From each paper, we extract the constraints on the context information model and its use for adaptive authentication systems put forward.

Metric. We analyse the properties and identify some that are commonly put forward.
The metric extracts the properties put forward in the reviewed publications and the frequency of papers putting them forward. We also analyse which of the properties are used as empirical evaluation metrics.

7.2 Findings on Desired Properties of the Context Information Model and its Use for Adaptive Authentication Systems

We extracted ten desired properties of the context model. Seven properties relate to the ability of the context model to handle specific characteristics of context information (1). The other three properties relate to the ability to be integrated in an adaptive authentication system (2).

(1) Properties related to the ability of the context model to handle specific characteristics of context information
- **Dynamicity:** The context model can take into account changes in the context information along the authentication process.
- **Quality:** The context model can evaluate the exactitude of the context information.
- **Temporality:** The context model can take into account temporal information which may impact the interpretation of the context.
- **Complexity:** The context model can consider the context as a mesh consisting of many different and connected information.
- **Heterogeneity,** The context model can take into account that the context consists of dissimilar or diverse information.
- **Abstraction:** The context model can reduce the amount of complexity of the context information.
- **Privacy:** The privacy requirements associated with the context information are taken into account in the model.

(2) Properties related to the ability of the context model to be integrated in an adaptive authentication system
- **System relevance:** The context model can provide machine interpretability and sufficient support for the authentication system's development process.
- **Accuracy:** The context model can reason about the context information in an accurate manner.

	Dynamicity	Quality	Temporality	Complexity	Heterogeneity	Abstraction	Privacy	System Relevance	Accuracy	Response Time
Al-Muhtadi et al. (2011) [3]	•		•	•	•		•	•		
Liu et al. (2021) [30]		•	•						••	
Kumar et al. (2021) [26]		•		•					••	
Solano et al. (2020) [52]	•		•	•	•		•		••	
Pititheeraphab et al. (2020) [44]				•				•	••	
Gunjal et al. (2020) [18]	•		•				•			
Miraoui et al. (2019) [36]						•			•	
Ma et al. (2018) [31]								•	••	••
Mozzaquatro et al. (2017) [37]	•	•	•		•	•				
Arias-Cabarcos et al. (2017) [4]	•		•	•	•		•			
El-Tarhouni et al. (2017) [15]		•		•		•	•		••	
Kumar et al. (2017) [27]	•	•	•						••	•
Neverova et al. (2016) [38]	•	•	•	•			•	•	••	•
Milton et al. (2016) [34]	•		•						••	
Ramakrishnan et al.(2015) [46]	•		•	•			•	••	••	
Perumal et al. (2015) [40]			•	•					••	••
Roth et al. (2014) [48]	•		•						••	•
Samyama et al. (2014) [50]	•		•							••
Witte et al. (2013) [57]	•								••	••
Cai et al. (2012) [10]		•					•		••	
Kisku et al. (2012) [25]		•	•	•		•			••	
En-Nasry et al. (2011) [16]	•		•	•			•	•		
Saedi et al. (2011) [49]			•	•					••	
Lima et al. (2011) [29]	•	•	•	•					••	

Table 6. Overview: Addressed[18] Desired Properties of the Context Information Model and its Use for Adaptive Authentication Systems

- **Response time:** The context model can reduce the total amount of time it takes to respond to an authentication request.

Table 6 shows an overview of which authors of the publications relevant to this article put forward which desired properties. A bullet means that the authors put forward the property in the discussion of their approach. Two bullets mean that the authors use the property as an empirical evaluation metric.

Dynamicity (58%). In some works the dynamicity of the users' behaviour is taken into account in the context model [18, 34, 38, 48, 52, 57]. Other authors model context in highly dynamic environments [4, 29, 37, 50]. Kumar et al. [27] study phone movement patterns under static and dynamic conditions. Ramakrishan et al. [46] assume security politics to be dynamic. The authentication of mobile dynamic identities is addressed in [16] and [3].

Quality (38%). Some authors analyse the quality of contextual information [10, 15, 25, 26, 30, 37]. The quality of classification algorithms for the classification of context information is discussed in some works [27, 38]. Lima et al. [29] analyse the quality of sensors to acquire context information.

Temporality (71%). Some authors analyse the temporal dimension of contextual features (*e.g.,* the hour of the connection) [3, 4, 16, 25, 30, 40, 46, 48, 52]. To take into account the *temporal* dimension, Gunjal et al. [18] propose checking the users' credentials on a *periodic* basis. In some works, the challenge of providing anytime authentication services, e.g. in ubiquitous systems [50] or the *Internet of Things* (IoT) [37], is discussed. In [29], the used space-time permutation model allows to take into account the temporal dimension of contextual features. The contextual features are analysed in different time windows in [27] and [57]. The use of *time* series data in [48, 49], enables taking into account the temporal dimension of contextual information.

Complexity (54%). Kumar et al. [26] discuss the complexity that human beings have almost the same palmprints. The complexity of the users' behaviour is discussed in some works [29, 52]. Pititheeraphab et al. [44] discuss the complexity of image processing for the representation of context information. The complexity of algorithms to reason about context

information is discussed in various works [4, 15, 38, 46]. In [25, 40, 49], the complexity of patterns is taken into account. The complexity of mobile identities is discussed in [16]. Al-Muhtadi et al. [3] model the complex usage patterns of devices in IoT environments and hence address the complexity of the contextual feature.

Heterogeneity (17%). Access patterns are assumed to be heterogeneous (*e.g.,* connections from multiple devices and locations due to travel) in [52]. Mozzaquatro et al. [37] discuss business opportunities based on a heterogeneous network of objects and their owners over the internet. Arias-Carbacos et al. [4] discuss the heterogeneity of authentication mechanisms in different contexts. In [3], the heterogeneity of IoT devices is discussed.

Abstraction (17%). To take into account the condition of reducing the amount of complexity, Miraoui et al. [36] discuss the right abstraction level of context to reduce and limit the set of contextual information. Multiple abstraction levels to provide meaningful information to understand the environment are discussed in [37]. In [15], the palmprints are represented on an abstracted level. Different abstraction levels of image fusion schemes are discussed in [25].

Privacy (38%). Several works address privacy issues related to context modelling. To take into account the condition of protecting private information, Solano et al. [52] split the keyboard in different areas to reduce privacy concerns for the analysis of keystrokes. Unacceptable privacy invasion is discussed in [18]. *Privacy* issues concerning the collection of user data are discussed in [4], [15] and [16]. Neverova et al. [38] discuss privacy issues concerning cloud computing. The users' needs regarding the protection of private data in social media is discussed in [46]. Private keys are used for the embedding algorithm in [10]. Al-Muhtadi et al. [3] aim for privacy protection with the help of third parties (clouds). We observe that privacy is still rather abstract and there is no clear consensus in the field of authentication on which data belongs to the user and which data can be exploited.

System Relevance (25%). To take into account the condition of providing machine interpretability and sufficient support for the system's development process, authors aim to ensure the ease of implementation [16, 44]. In [31], the processing power of the central server is taken into account. The storage, memory and processing power of devices is addressed in [38]. The system relevance is evaluated empirically in [46] in terms of energy efficiency. Al-Muhtadi et al.'s [3] framework is implemented in the IBM cloud platform.

Accuracy (75%). Many authors calculate accuracy metrics (*e.g., Equal Error Rate* (EER), *False Positive Rate* (FPR), *False Negative Rate* (FNR)) to evaluate their approaches [10, 15, 25–27, 29–31, 34, 38, 40, 44, 46, 48, 48, 49, 52, 57].

Response Time (29%). To take into account the amount of time it takes to respond to a request for a service, several authors discuss the speed of their algorithms [27, 38]. Metrics for evaluating the response time of the system are proposed in [31, 40, 57]. Roth et al.'s [48] overall goal is to explore a biometric with short response time for detection. Samyama et al. [50] evaluate empirically the time spend for the generation of authentication certificates.

Successful context models for adaptive authentication systems have at least some of these properties, although almost no context models have them all. As CM4AA is a cross-cutting concern in multiple domains, there is a great diversity of desired properties, which play different roles in the different domains. Also, the right balance between the properties varies from domain to domain. *Accuracy*, which is the ability of the context model to reason about the context information in an accurate manner, is put forward in 75% of the reviewed papers. Biometrics are frequently used contextual features and biometric system accuracy testing is common. Also, we have seen that it is common to use contextual features that clearly determine an unique identity. The *accuracy* of such determinations is crucial. In almost every work

(94%) which is addressing *accuracy*, the property is evaluated empirically with the help of common metrics (*e.g.,* FPR, EER). These are metrics often used to evaluate the performance of machine learning algorithms. For CM4AA, it is common to use learning algorithms, for example to detect derivations from patterns or other anomalies. Often, their *accuracy* is evaluated. The properties *response time* and *system relevance* are evaluated empirically in some works as well. Overall, however, only one third of the properties are evaluated empirically. The desired properties of the context model seem not to be standardised enough (*e.g.,* there are no benchmark solutions for how to take into account changes in the context information along the authentication process), what is also due to the fact that needs vary greatly across the different application domains. Another frequently addressed property is temporality (71%). It is common to take into account the temporal dimension of contextual information which may change its interpretation. Patterns and user habits are often based on time. The ability to take into account the changes in the context information along the authentication process is addressed as desired property in 58% of the reviewed works. The authors consider aspects of the environment that may change in the authentication system.

Lessons Learned. We observe a great diversity of desired properties of the context information model and its use for adaptive authentication systems due to the fact that CM4AA is a cross-cutting concern in multiple domains. The ten observed desired properties can be divided into two classes: properties related to the ability of the context model to handle specific characteristics of context information (1), and properties related to the ability of the context model to be integrated in an adaptive authentication system (2). Successful context models for adaptive authentication systems have at least some of these properties, although almost no context models have them all. A big challenge is to find the right balance between different properties. Very commonly the properties accuracy (75%), temporality (71%) and dynamicity (58%) are put forward. To evaluate the properties empirically benchmark solutions are missing.

8 SWOT MATRIX - *(STRENGTHS, WEAKNESSES, OPPORTUNITIES, THREATS)*

We summarise our findings in a SWOT analysis on CM4AA. SWOT analysis is a technique for assessing strengths, weaknesses, opportunities, and threats. With this tool, we aim to analyse what is done best right now in the research area of CM4AA, and to devise a successful strategy for future research and practice. Fig. 12 shows the SWOT Matrix, which we derive from our analysis.

Strengths. Strengths are things that are done particularly well in the research area of CM4AA. Research conducted by observing and analyzing context information for adaptive authentication systems and resulting in abstract **concepts and ideas** is well advanced. The ability of (mobile) devices to sense their physical environment and adapt their behavior accordingly (context-awareness) is helpful to successfully model context for adaptive authentication systems. Another strength is the capability to analyse **biometric and behavioral information**. These also exist thanks to modern technologies and advancements in the research area. Also, accurate approaches for **anomaly detection** exist to detect derivations from patterns.

Weaknesses. Harmful to successfully model context information for adaptive authentication systems is the **lack of standards and benchmark solutions**, which makes it difficult to compare approaches or to present a holistic overview of context information for adaptive authentication systems. **Public data** is missing, and companies do not publish their **state of the practice**. There are only **few tools** for modelling context information for adaptive authentication systems what makes it difficult for adaptive authentication system designers to use context information efficiently. There are only few works treating context CM4AA at **runtime**. The **context of other entities than the user** is

often disregarded. There are many works focusing on a limited set of contextual features, but there is a lack of works regarding what context information can be used for adaptive authentication in a **holistic manner**.

Opportunities. Despite the weaknesses, there is a great variety of opportunities in the research field of CM4AA. There are more and more opportunities for **context awareness** thanks to the ability of (mobile) devices to sense their physical environment and adapt their behavior accordingly. CM4AA is a **young research area** and we observe a **steady interest** in the topic. **Mobile computing environments** are great opportunities, especially for data acquirement. Another opportunity is the **use of less privacy-sensitive context information** in cases in which it is not necessary to identify a unique entity. **Privacy regulation standards** like *General Data Protection Regulation* (GDPR) can also be seen as an opportunity for the research area. Having different restrictions in different countries extend the scope of adaptability. Having guidelines allows adapting in a regulated manner. Also, **anomalies** that are not based on the user's patterns and habits are an opportunity in the research area.

Threats. We also identify threats harming successful CM4AA. The GDPR data protection standard is a threat regarding **private data collection**. It can be difficult to acquire contextual information according to these restrictions. **Disregarding non-mobile devices** is a threat as well. Often, approaches are based on mobile devices and their sensing abilities. If adaptive authentication is used on non-mobile devices, the data must be acquired differently. For example, the contextual feature "location" can be acquired easily from mobile devices equipped with GPS sensors, but hardly from non-mobile devices. The **interdisciplinary** of the research area is a threat as well because notions and needs differ across the disciplines. We have seen that the balance between desired properties of authentication mechanisms is crucial for adaptive authentication. This balance may also depend on the domain. The **heterogeneity of context information and devices** is another important threat because they need to be taken into account when the context information is modelled for adaptive authentication systems. Desired properties of the context information model and its use for adaptive authentication systems are still rather abstract and it is hard to evaluate them empirically.

9 THREATS TO VALIDITY OF OUR STUDY

Troya et al. [54] study four basic types of validity threats that can affect studies like ours. We cover three of them in the following. As our work is a review of a specific topic, we do not intend to make any generalizations and hence do not treat the threat type *external validity*.

Conclusion validity. Issues that affect the ability to draw conclusions and whether the survey can be repeated concern the conclusion validity [54]. The availability of the raw search results and the set of excluded studies on our website mitigates these threats. Our analysis metrics can easily be repeated and verified. Like Troya et al. [54] we did not include works not (yet) published or submitted even if they might alter the results of our study. We assume that the disadvantages of inclusion (*e.g.,* lack of quality, difficulty of identification) outweigh the advantages. We are aware that the number of our articles is relatively small. As there are many different works in the field of context-awareness and modelling, we prefer to concentrate on this particular selection of works to ensure the meaningfullness of our analysis for authentication systems.

Construct validity. We mitigate the issue known as meno-method bias [54], that might arise during research design by following the methodologies of SMS and SLR. Another threat regarding the construct validity is that particular works can be categorised in more than one dimension of our analysis aspects. We mitigate this issue by assigning the dimension that fits best according to multiple analysts from the authentication and the modelling domain. We observe

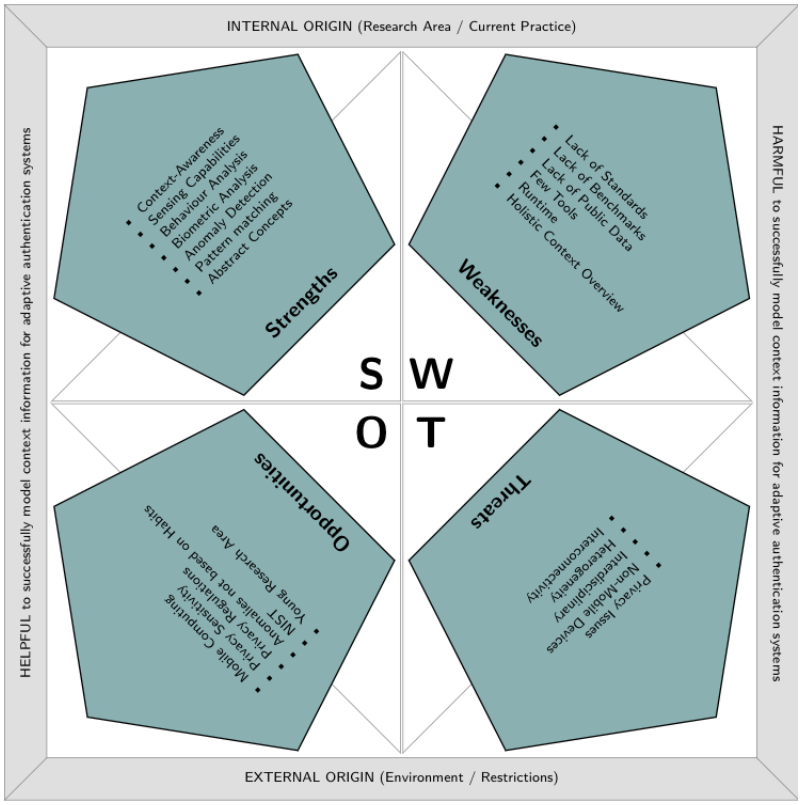

Fig. 12. Research field of CM4AA - SWOT Matrix

that there is no clear consensus of which are the most important properties of the context information model and its use for adaptive authentication systems. The definition of the terms is still rather abstract. Our analysis therefore only gives an indication of what can be crucial, but we do not have any evidence to justify that if none of these properties is satisfied, the technique is not successful.

Internal validity. According to [54] the main factors influencing the publication selection process and therefore affecting the results of our evaluation are keywords, digital libraries, the language of publication, and time frame. We avoid too restrictive decisions by including a disjunction of terms expressing the adaptation capability of the authentication system in our search clause. Also, we included different spellings of the terms. To mitigate sampling and publication bias, we conduct searches on formal databases (*e.g.,* ACM Digital Library) and indexes (*e.g.,* GoogleScholar).

10 RELATED SURVEYS

Most of the surveys related to our work fall in the categories of *context and context-awareness*, *authentication modalities*, *adaptive authentication in specific computing environments* and *adaptive authentication in general*. In the following, we present existing reviews and studies belonging to these topics.

Context and Context-awareness. Works in the literature studied context, and context-awareness. Habib and Leister [19] present the concepts of context, context-awareness and context-based security. They present an overview of context-awareness definitions and explain the life-cycle process of a context-aware system. The work includes summaries of context types, context attributes, and context modelling approaches. Other reviews focus on **co-presence detection** and **proximity sensing** for determining contexts. Contextual **co-presence detection** is focused in the work of Truong and Asokan [51]. Shrestha et al. [51] investigate **sensor-based fusion approaches** for **proximity detection** of devices and nearby people in the face of active adversaries. A study [55] shows the potential of **fusing** multiple sensor modalities for better resilience against certain attack types. The authors investigate the use of different **co-presence detection** sensors and their fusions.

Authentication Modalities. Other related literature reviews extensively analyse specific modalities for authentication. Mir et al. [35] propose a literature survey on **biometrics** verification. Baldini and Steri [7] analyse techniques using physical **fingerprints**. Existing **eye movement** authentication methods are reviewed comparatively by Das et al. [11]. Pisani et al. [43] review adaptive approaches for **keystroke dynamics**. They outline the need for models that adapt dynamically to changes in users' typing behaviours. Algorithms for user authentication based on **keystroke dynamics** are evaluated in this work, and several modifications are proposed for making them able to dynamically adapt their behaviour in time.

Adaptive authentication in specific computing environments. Other surveys consider adaptive authentication in specific **computing environments**. Kayes et al. [23] propose a review of the current literature in the field of context-aware access control for **cloud and fog computing**. Stojanov et al. [53] propose a ranking of existing **semantic web autorisation systems**. Khan et al. [24] review trust management techniques in the **social internet of things**. Context-aware authentication for the **IoT** is focused in the work from Habib and Leister [19]. Pal et al. [39] outline classifications and trends for identity modelling for the **IoT**. A study on access control approaches in the context of **IoT** is proposed by Al-Halabi et al. [2]. The authors aim to help researchers to define new models and systems for access control regarding new challenges due to the IoT environment.

Adaptive authentication in general. Our work is complementary to the survey on adaptive authentication from Arias-Cabarcos et al. [5]. In their work, the authors establish a common definition of adaptive authentication system, analyse adaptive authentication approaches and identify research challenges. The focus of their work is on how design principles well known in adaptive systems, can be applied on adaptive authentication systems. They provide an overview of "adaptation reasons" consisting of a set of contextual features describing the security context, the usability context, technical resources and the user and determine which changes in features lead to the need to adapt the system. In our work, we deeply study the **modelling of these contextual features.** We are interested in **how the context for adaptive authentication systems is modelled** and **how the context information model is used for adaptive authentications systems**. Arias-Cabarcos et al. [5] define an adaptive authentication system as a system that *"is able to automatically modify its behavior and/or structure in response to changes in its operating environment"*. We define an adaptive authentication system as *a context-aware authentication system that uses context to provide the relevant authentication mechanism(s), where relevancy depends on the desired properties of the authentication mechanism for a user in a context.* According to their definition, Arias-Cabarcos et al. [5] study how authentication systems adapt in response to changes in the context. They are interested in the adaptation logic of the system and consider authenticators as the elements that need to be adapted and discuss their properties. Hence, they do not limit the search space to

articles that explicitly contain "context modelling". They also consider papers that only contain "authentication" and not necessarily "authentication system". Questions about the adaptation logic can be answered with the help of such papers, but we can't get any information about **context information gathering, modelling, data structures, and their evaluation for authentication systems**. In our work, we aim to analyse how context information modelling for adaptive authentication systems is performed to analyse **how context models that are suitable** for the field are defined and evaluated. Complementary to [5] and leveraging on their conclusions, we aim (1) to find out whether there are standard means for context modeling given the gathering and availability constraints, (2) to uncover the desired properties of the context information models for adaptive authentication systems and (3) to analyse the properties enabling interoperability within adaptive systems that include different sensors, devices and platforms. We analyse the properties which enable their interoperability within adaptive systems that will potentially include different **sensors, devices and platforms**. Arias-Cabarcos et al. [5] outline that context modelling for security applications (e.g., adaptive authentication) has not been deeply studied until now, that the works surveyed in their article show a limited usage of context, with vague descriptions and grounds and that it is difficult to reuse or extend adaptive authentication systems due the lack of practical solutions. Within this work, we conduct efforts to find out **what models are suitable** for the field of context modelling for adaptive authentication. Our study is an important first step towards **less vague descriptions and grounds** of using context for authentication systems. In this work, we demonstrate the **ability of capturing a common set of contextual features** that are relevant for adaptive authentication systems independent from the application and show that despite the possibility of a unified framework, **no standard exists**. Our results are a first step towards more reusable and extendable adaptive authentication systems.

11 CONCLUSION AND PERSPECTIVES

Within this article, we identify the current body of knowledge about CM4AA, what context information determines the context of adaptive authentication systems, how the context information is modelled, how the context information model is used, and what are the desired properties of the context information model and its use for adaptive authentication. We shed light on three research questions and we offer an overview of existing research that security practitioners and non-domain experts can use. For each research question, we collected a certain amount of raw data on the selected articles, and we defined a set of metrics allowing us to analyse this raw data.

We observe a **continuous interest** in the research field of CM4AA over the last ten years. Most of the reviewed publications (91%) are **not specific to any application domain**. 16% of the contributions are of the contribution type **tool**. Adaptive authentication is a new research area, so that not yet every proposed concept of how to model context information for adaptive authentication systems goes beyond conceptualization and results in a tool. In the research field of CM4AA, it is widespread to acquire context information from **sensors of mobile devices** to describe the context of a **user**. The most frequently used contextual features for adaptive authentication systems are **biometrics**, the **entities behaviour** and the **location**. The contextual features are mostly analysed in **time**. We can not observe a trend in the use of a **modelling technique** to model context information for adaptive authentication systems but we can identify a set of common goals. There is a great diversity of modelling formalisms proposed in the literature. The context information models are mostly used at the **design time** (63%) and **deployment time** (42%) of adaptive authentication systems. There is a lack of works treating CM4AA at **runtime** (8%). According to the percentage of works putting forward each of the desired properties, accuracy (78%), temporality (74%), security(70%), and dynamicity (61%) seem to be the most important desired properties of the context information model and its use for adaptive authentication systems.

The great diversity regarding the choice of the context information, and the modelling approaches, makes it challenging to propose a one fits all solution for CM4AA. Anyway, practitioners need support regarding the **conception of context information models**. There is a need for a **modelling framework** for context modelling for adaptive authentication systems, which focuses on a holistic overview of context information for adaptive authentication systems. Adaptive authentication practitioners need to get **recommendations** regarding the use of context information for adaptive authentication systems. In the future, we plan to provide a **model-based framework** for context modelling for adaptive authentication systems. Within this framework, we plan to cover a maximum of aspects relevant to context modelling for adaptive authentication systems outlined in this article. We aim to provide a **recommendation tool**, which can be used to get support for modelling context information for adaptive authentication systems.

In this work, we focus on context modelling for adaptive authentication systems and do not discuss self-adaptive systems design in general. We conduct efforts to find out what models are suitable for the field. However, our results may be helpful for further research on adaptive system design in general.

Future reflections also need to be made regarding the **heterogeneity of mobile and non-mobile devices** and how adaptive authentication can work in both cases. Issues related to mobile computing and IoT environments, as the acquirement of context information, the multiplicity of devices, and privacy aspects, need to be treated.

The focus of future work also needs to be on pushing further the **implementation** of concepts and model designs.

Another interesting aspect for future research is the question of how to gather **benchmark solutions** for context modelling for adaptive authentication systems and **public data** for evaluation.

REFERENCES

[1] Achilleas P. Achilleos, Georgia M. Kapitsaki, and George A. Papadopoulos. 2012. A Framework for Dynamic Validation of Context-Aware Applications. In *2012 IEEE 15th International Conference on Computational Science and Engineering*. 532–539. https://doi.org/10.1109/ICCSE.2012.79

[2] Yahia Al-Halabi, Nisreen Raeq, and Farah Abu-Dabaseh. 2017. Study on Access Control Approaches in the Context of Internet of Things: A survey. In *2017 International Conference on Engineering and Technology (ICET)*. IEEE, Akdeniz University, Turkey, 1–7.

[3] Jalal Al-Muhtadi, Kashif Saleem, Sumayah Al-Rabiaah, Muhammad Imran, Amjad Gawanmeh, and Joel JPC Rodrigues. 2021. A lightweight cyber security framework with context-awareness for pervasive computing environments. *Sustainable Cities and Society* 66 (2021), 102610.

[4] Patricia Arias-Cabarcos and Christian Krupitzer. 2017. On the Design of Distributed Adaptive Authentication Systems. *Open Access Media* 5 (2017), 12–14.

[5] Patricia Arias-Cabarcos, Christian Krupitzer, and Christian Becker. 2019. A survey on Adaptive Authentication. *ACM Computing Surveys (CSUR)* 52, 4 (2019), 1–30.

[6] Khairul Azmi Abu Bakar and Galoh Rashidah Haron. 2013. Adaptive Authentication: Issues and challenges. In *2013 World Congress on Computer and Information Technology (WCCIT)*. IEEE, Dhaka, Bangladesh, 1–6.

[7] Gianmarco Baldini and Gary Steri. 2017. A survey of Techniques for the Identification of Mobile Phones Using the Physical Fingerprints of the Built-in Components. *IEEE Communications Surveys & Tutorials* 19, 3 (2017), 1761–1789.

[8] Nelly Bencomo, Sebastian Götz, and Hui Song. 2019. Models@ run. time: a guided tour of the state of the art and research challenges. *Software & Systems Modeling* 18, 5 (2019), 3049–3082.

[9] Emmanuel Bertin, Dina Hussein, Cigdem Sengul, and Vincent Frey. 2019. Access control in the Internet of Things: a survey of existing approaches and open research questions. *Annals of Telecommunications* 74 (03 2019). https://doi.org/10.1007/s12243-019-00709-7

[10] Li-jun Cai, Rui Li, and Ye-qing Yi. 2012. A multiple watermarks algorithm for image content authentication. *Journal of Central South University* 19, 10 (2012), 2866–2874.

[11] I. Das, S. Singh, R. Das, S. Biswas, S. Roy, and S. Gupta. 2020. Design and Implementation on EMBA Authentication models. In *2020 IEEE VLSI DEVICE CIRCUIT AND SYSTEM (VLSI DCS)*. IEEE, Kolkata, India, 283–288. https://doi.org/10.1109/VLSIDCS47293.2020.9179890

[12] Anind K Dey. 2001. Understanding and Using Context. *Personal and ubiquitous computing* 5, 1 (2001), 4–7.

[13] Ana I Segovia Domingo and Álvaro Martín Enríquez. 2018. Digital Identity: the current state of affairs. *BBVA Research* 1, 0 (2018), 1–46.

[14] Claudia Eckert. 2013. *IT-Sicherheit: Konzepte-Verfahren-Protokolle*. Walter de Gruyter, Germany.

[15] Wafa El-Tarhouni. 2017. *Finger Knuckle Print and Palmprint for Efficient Person Recognition*. Ph.D. Dissertation. Northumbria University, Northumbria.

[16] Brahim En-Nasry and Mohamed Dafir Ech-Cherif El Kettani. 2011. Towards an open framework for mobile digital identity management through strong authentication methods. In *FTRA International Conference on Secure and Trust Computing, Data Management, and Application*. Springer, na,

56–63.

[17] David Freeman, Sakshi Jain, Markus Dürmuth, Battista Biggio, and Giorgio Giacinto. 2016. Who Are You? A Statistical Approach to Measuring User Authenticity.. In *NDSS*, Vol. 16. 21–24.

[18] Samyama Gunjal GH and Samarth C Swamy. 2020. A Security Approach to Build a Trustworthy Ubiquitous Learning System. In *2020 IEEE Bangalore Humanitarian Technology Conference (B-HTC)*. IEEE, Karnataka, India, 1–6.

[19] Kashif Habib and Wolfgang Leister. 2015. Context-Aware Authentication for the Internet of Things. In *The Eleventh International Conference on Autonomic and Autonomous Systems*. IEEE, Rome, Italy, 134–139.

[20] Daniel Hintze, Matthias Füller, Sebastian Scholz, Rainhard D Findling, Muhammad Muaaz, Philipp Kapfer, Eckhard Koch, and René Mayrhofer. 2019. CORMORANT: Ubiquitous risk-aware multi-modal biometric authentication across mobile devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 1–23.

[21] Syed Zulkarnain Syed Idrus, Estelle Cherrier, Christophe Rosenberger, and Jean-Jacques Schwartzmann. 2013. A review on authentication methods. *Australian Journal of Basic and Applied Sciences* 7, 5 (2013), 95–107.

[22] Gleneesha M Johnson. 2009. Towards shrink-wrapped security: A taxonomy of security-relevant context. In *2009 IEEE International Conference on Pervasive Computing and Communications*. IEEE, 1–2.

[23] ASM Kayes, Rudri Kalaria, Iqbal H Sarker, Md Islam, Paul A Watters, Alex Ng, Mohammad Hammoudeh, Shahriar Badsha, Indika Kumara, et al. 2020. A Survey of Context-Aware Access Control Mechanisms for Cloud and Fog Networks: Taxonomy and Open Research Issues. *Sensors* 20, 9 (2020), 2464.

[24] Wazir Zada Khan, Saqib Hakak, Muhammad Khurram Khan, et al. 2020. Trust Management in Social Internet of Things: Architectures, Recent Advancements and Future Challenges. *IEEE Internet of Things Journal* 8, 10 (2020), 7768–7788.

[25] Dakshina Ranjan Kisku, Ajita Rattani, Phalguni Gupta, Jamuna Kanta Sing, and C Jinshong Hwang. 2012. Human Identity Verification Using Multispectral Palmprint Fusion. *Journal of Signal and Information Processing* 3, 2 (2012), 263–273.

[26] Abhilove Kumar and Apoorv Mishra. 2021. Palm print Recognition using 2D Fourier Transformation and Integration Function. *na* na, na (2021), na.

[27] Rajesh Kumar, Partha Pratim Kundu, Diksha Shukla, and Vir V Phoha. 2017. Continuous User Authentication via Unlabeled Phone Movement Patterns. In *2017 IEEE International Joint Conference on Biometrics (IJCB)*. IEEE, Denver, CO, USA, 177–184.

[28] Nilesh A Lal, Salendra Prasad, and Mohammed Farik. 2016. A review of authentication methods. *vol* 5 (2016), 246–249.

[29] Joao Carlos D Lima, Cristiano C Rocha, Iara Augustin, et al. 2011. A Context-Aware Recommendation System to Behavioral Based Authentication in Mobile and Pervasive Environments. In *2011 IFIP 9th International Conference on Embedded and Ubiquitous Computing*. IEEE, Melbourne, Australia, 312–319.

[30] Meng Liu, Longbiao Wang, Kong Aik Lee, Hanyi Zhang, Chang Zeng, and Jianwu Dang. 2021. Exploring Deep Learning for Joint Audio-Visual Lip Biometrics.

[31] Sihua Ma et al. 2018. *Using Blockchain to Build Decentralized Access Control in a Peer-to-Peer E-Learning Platform*. Ph.D. Dissertation. University of Saskatchewan, Saskatchewan.

[32] Alfred J Menezes, Paul C Van Oorschot, and Scott A Vanstone. 2018. *Handbook of applied cryptography*. CRC press, na.

[33] Markus Miettinen, Thien Duc Nguyen, Ahmad-Reza Sadeghi, and N Asokan. 2018. Revisiting context-based authentication in IoT. In *Proceedings of the 55th Annual Design Automation Conference*. 1–6.

[34] Leslie C Milton and Atif Memon. 2016. Intruder Detector: A Continuous Authentication Tool to Model User Behavior. In *2016 IEEE Conference on Intelligence and Security Informatics (ISI)*. IEEE, Tucson, AZ, USA, 286–291.

[35] AH Mir, S Rubab, and ZA Jhat. 2011. Biometrics Verification: a Literature Survey. *International Journal of Computing and ICT Research* 5, 2 (2011), 67–80.

[36] Moeiz Miraoui and Sherif El-etriby. 2019. A Context-Aware Authentication Approach for Smartphones. In *2019 International Conference on Computer and Information Sciences (ICCIS)*. IEEE, Aljouf, Kingdom of Saudi Arabia, 1–5.

[37] Bruno A Mozzaquatro, Ricardo Jardim-Goncalves, and Carlos Agostinho. 2017. Situation Awareness in the Internet of Things. In *2017 International Conference on Engineering, Technology and Innovation (ICE/ITMC)*. IEEE, Madeira Island, Portugal, 982–990.

[38] Natalia Neverova, Christian Wolf, Griffin Lacey, Lex Fridman, Deepak Chandra, Brandon Barbello, and Graham Taylor. 2016. Learning Human Identity from Motion Patterns. *IEEE Access* 4 (2016), 1810–1820.

[39] S. Pal, M. Hitchens, and V. Varadharajan. 2018. Modeling Identity for the Internet of Things: Survey, Classification and Trends. In *2018 12th International Conference on Sensing Technology (ICST)*. ICST, Limerick, Ireland, 45–51. https://doi.org/10.1109/ICSensT.2018.8603595

[40] Esther Perumal and Shanmugalakshmi Ramachandran. 2015. A Multimodal Biometric System Based on Palmprint and Finger Knuckle Print Recognition Methods. *International Arab Journal of Information Technology (IAJIT)* 12, 2 (2015), 118–128.

[41] Kai Petersen, Robert Feldt, Shahid Mujtaba, and Michael Mattsson. 2008. Systematic mapping studies in software engineering. In *12th International Conference on Evaluation and Assessment in Software Engineering (EASE) 12*. na, na, 1–10.

[42] Kai Petersen, Sairam Vakkalanka, and Ludwik Kuzniarz. 2015. Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology* 64 (2015), 1–18.

[43] P.H. Pisani, A.C. Lorena, and A.C.P.L.F. de Carvalho. 2015. Adaptive Approaches for Keystroke Dynamics. In *2015 International Joint Conference on Neural Networks (IJCNN)*. IJCNN, Killarney, Ireland, 1–8. https://doi.org/10.1109/IJCNN.2015.7280467

[44] Yutthana Pititheeraphab, Nuntachai Thongpance, Hisayuki Aoyama, and Chuchart Pintavirooj. 2020. Vein Pattern Verification and Identification Based on Local Geometric Invariants Constructed from Minutia Points and Augmented with Barcoded Local Feature. *Applied Sciences* 10, 9 (2020), 3192.

[45] Martin F Porter. 1980. An algorithm for suffix stripping. *Program* 14, 3 (1980), 130–137.

[46] Arun Ramakrishnan, Jochen Tombal, Davy Preuveneers, and Yolande Berbers. 2015. PRISM: Policy-Driven Risk-Based Implicit Locking for Improving the security of Mobile End-User Devices. In *Proceedings of the 13th International Conference on Advances in Mobile Computing and Multimedia*. ACM, Brussels, Belgium, 365–374.

[47] Oriana Riva, Chuan Qin, Karin Strauss, and Dimitrios Lymberopoulos. 2012. Progressive authentication: deciding when to authenticate on mobile phones. In *21st USENIX Security Symposium (USENIX Security 12)*. 301–316.

[48] Joseph Roth, Xiaoming Liu, and Dimitris Metaxas. 2014. On Continuous User Authentication via Typing Behavior. *IEEE Transactions on Image Processing* 23, 10 (2014), 4611–4624.

[49] Shahla Saedi and Nasrollah Moghadam Charkari. 2011. Characterization of palmprint using discrete orthonormal s-transform. In *2011 International Conference on Hand-Based Biometrics*. IEEE, na, 1–6.

[50] GH Samyama Gunjal, Pallapa Venkataram, and G Narendra Kumar. 2014. A Context-Based User Authentication Scheme for Ubiquitous Services. In *Proceedings of the World Congress on Engineering and Computer Science*, Vol. 1.

[51] Babins Shrestha, Nitesh Saxena, Hien Thi Thu Truong, and N Asokan. 2018. Sensor-based Proximity Detection in the face of Active Adversaries. *IEEE Transactions on Mobile Computing* 18, 2 (2018), 444–457.

[52] Jesus Solano, Luis Camacho, Alejandro Correa, Claudio Deiro, Javier Vargas, and Martín Ochoa. 2019. Risk-based Static Authentication in Web Applications with Behavioral Biometrics and Session Context Analytics. In *International Conference on Applied Cryptography and Network Security*. Springer, Bogotá, Colombia, 3–23.

[53] Riste Stojanova, Slobodanka Stojanovab, Milos Jovanovika, Vladimir Zdraveskia, and Dimitar Trajanova. 2017. Ranking Semantic Web Authorization Systems. *Semantic Web 8* 1, 5 (2017), 570–0844.

[54] Javier Troya, Nathalie Moreno, Manuel F Bertoa, and Antonio Vallecillo. 2021. Uncertainty representation in software models: a survey. *Software and Systems Modeling* na, na (2021), 1–31.

[55] H. T. T. Truong, Xiang Gao, B. Shrestha, N. Saxena, N. Asokan, and P. Nurmi. 2014. Comparing and Fusing Different Sensor Modalities for Relay Attack Resistance in Zero-Interaction Authentication. In *2014 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, Budapest, Hungary, 163–171. https://doi.org/10.1109/PerCom.2014.6813957

[56] Stephan Wiefling, Luigi Lo Iacono, and Markus Dürmuth. 2019. Is this really you? An empirical study on risk-based authentication applied in the wild. In *IFIP International Conference on ICT Systems Security and Privacy Protection*. Springer, 134–148.

[57] Heiko Witte, Christian Rathgeb, and Christoph Busch. 2013. Context-Aware Mobile Biometric Authentication Based on Support Vector machines. In *2013 Fourth International Conference on Emerging Security Technologies*. IEEE, Cambridge, United Kingdom, 29–32.