2020 Second IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA) | 978-1-7281-8543-9/20/$31.00 ©2020 IEEE | DOI: 10.1109/TPS-ISA50397.2020.00034

AuthN-AuthZ: Integrated, User-Friendly and Privacy-Preserving Authentication and Authorization

Tyler Phillips*, Xiaoyuan Yu*, Brandon Haakenson*,
Shreya Goyal†, Xukai Zou*, Saptarshi Purkayastha†, Huanmei Wu†
*Department of Computer Science, †Department of BioHealth Informatics
Indiana University-Purdue University Indianapolis
Indianapolis, Indiana 46202, USA
{phillity, xyu1, bhaakens, shregoya, xzou, saptpurk, hw9}@iupui.edu

***Abstract*—In this paper, we propose a novel, privacy-preserving, and integrated authentication and authorization scheme (dubbed as AuthN-AuthZ). The proposed scheme can address both the usability and privacy issues often posed by authentication through use of privacy-preserving Biometric-Capsule-based authentication. Each Biometric-Capsule encapsulates a user's biometric template as well as their role within a hierarchical Role-based Access Control model. As a result, AuthN-AuthZ provides novel efficiency by performing both authentication and authorization simultaneously in a single operation. To the best of our knowledge, our scheme's integrated AuthN-AuthZ operation is the first of its kind. The proposed scheme is flexible in design and allows for the secure use of robust deep learning techniques, such as the recently proposed and current state-of-the-art facial feature representation method, ArcFace. We conduct extensive experiments to demonstrate the robust performance of the proposed scheme and its AuthN-AuthZ operation.**

***Index Terms*—Authentication, Authorization, Face Recognition, Deep Learning, Biometric-Capsule, AuthN-AuthZ**

I. Introduction

With the recent proliferation of biometric sensors in smart devices, biometrics-based authentication has emerged as a promising solution to address the inherent usability issues associated with traditional information and object-based authentication methods [1]. Unfortunately, the acceptance of biometrics-based authentication is challenged by several factors. A recent and stark shift in attitudes has led many users who first perceived emerging biometrics-based technologies as usable, secure and attractive, to view biometrics with apprehension and skepticism [2]. User privacy concerns surrounding robust, recently proposed deep learning-based biometrics techniques have overshadowed the usability benefits of biometrics. As a result, the public has begun to perceive biometrics as a tool to violate privacy and to facilitate oppression [3]. In the United States, these concerns have culminated in calls for regulations to be imposed on biometrics-based technologies [4].

To address these pressing privacy issues, many secure biometric authentication methods have been proposed [5], [6]. One such method is the fusion-based Biometric-Capsule, or BioCapsule, (BC) scheme [7], [8]. The BC method secures biometric authentication systems by fusing sampled user biometrics with the biometrics of a reference subject (RS) (see Section III-A for more discussion). The BC generation process combines several one-way functions which serve to mask the contributions of user and RS biometrics within their resulting BC (see Figure 1). As a result, the BC scheme is provably secure, privacy-preserving and robust against several types of attacks [7].

Authorization, typically carried out after authentication, is another essential process in computing systems in which a user is granted data and operating privileges. Many access control models have been proposed in order to allow system administrators to exercise fine-tuned control over user privileges [9]. Role-Based Access Control (RBAC) models [10] are widely adopted and allow system administrators to model complex, hierarchical relationships between different roles. RBAC models first group users based on their roles within a computing system. Then, each role is granted corresponding privileges for data and operating access.

Currently, most computing systems implement authorization and authentication independently. In many proposed systems, there is typically an implicit assumption that authorization will be simply addressed after successful authentications. Unfortunately, assumptions such as this have led to systems with weak and insufficient access control mechanisms [11].

When authentication and authorization are loosely coupled it can lead to vulnerabilities due to a mismatch of what information each part of the system expects. Generally, it is expected that the system will have verified a user's identity before they reach a stage where an authorization check is required. However, it may be possible for an attacker to take advantage of the gaps between these steps. In some cases, a user may be able to pass the authentication step normally but then spoof their identity as that of another user before the authorization step. The authorization step may then implement an access control check using the fake identity rather than the one that was authenticated. In a loosely coupled system, a user's identity may not be directly tied to their permissions.

Instead, authorization could be directly addressed by tightly coupling both authentication and authorization into a single integrated operation. It is a challenging task, which has not been addressed by previous research. We propose a novel, combined authentication and authorization operation dubbed AuthN-AuthZ. The proposed operation is facilitated using a state-of-the-art and deep learning-enhanced BC-embedded

978-1-7281-8543-9/20/$31.00 ©2020 IEEE
DOI 10.1109/TPS-ISA50397.2020.00034

facial authentication scheme (outlined in Section V). During BC-embedded facial authentication, a user's RS will denote the user's role within a RBAC model [10] (as explained in Section IV). Therefore, a user's BC will securely encapsulate both a user's biometric features (authentication credentials) and the user's RBAC role (access control privilege). As a result, a user will be both authenticated and authorized in a single AuthN-AuthZ operation. To the best of our knowledge, the proposed integrated AuthN-AuthZ operation is the first of its kind. This work provides the following contributions:

1) The BC-embedded facial AuthN-AuthZ system is able to address (i) the usability issues associated with knowledge and object-based authentication and (ii) the inherent privacy and revocability issues of unsecured biometrics-based authentication.
2) The BC-embedded facial AuthN-AuthZ system is able to resolve the general performance issues often posed by secure biometric authentication by leveraging recently proposed state-of-the-art deep learning-based techniques.
3) The innovative, user-friendly and flexible AuthN-AuthZ operation is able to directly resolve the separation, inefficiency and inconvenience of handling authentication and authorization independently.
4) We openly provide all code necessary to implement the AuthN-AuthZ system and replicate our experimental results [12].

II. Related Work

Although biometrics-based authentication addresses usability issues that are inherent within knowledge and object-based authentication methods, it also introduces pressing privacy concerns [5]. One major drawback of biometric authentication is that a user cannot revoke or replace their biometric traits (as they could for a password or smartcard). If a user's biometric template is stolen by an attacker, it is forever compromised. An attacker can use a stolen biometric template for a variety of attacks such as masquerading attacks, replay attacks, spoofing attacks and cross-application attacks [5]. Furthermore, analysis of stolen biometric templates can also reveal sensitive personal information about a victim user, such as age, ethnicity, and gender [13]. How to effectively secure the biometric templates used within authentication systems has been an active area of research for many years [5], [6]. Two main classes of approaches have emerged: Biometric Cryptosystems (BCS) and Cancellable Biometrics (CB).

BCS approaches generate authentication keys from sampled biometrics, instead of using sampled biometrics directly [6]. There are two main types of schemes within BCS: key generating schemes [14] and key binding schemes [15]. CB methods transform sampled biometrics and use the resulting altered biometric templates for authentication [5]. CB schemes can be grouped into two main types: biometric salting [16] and noninvertible transformations [7].

Unfortunately, BCS and CB schemes are not often provably secure and suffer from many types of attacks [6]. In addition, CB and BCS schemes commonly have significant adverse effects on authentication performance [5]. This is typically because many CB and BCS schemes present a trade-off between authentication performance and biometric template privacy. Furthermore, many proposed schemes are not interoperable, i.e., they require fixed preprocessing, feature extraction or classification techniques. Unfortunately, these required techniques are often dated machine learning and computer vision algorithms that can not offer robust authentication performance. The resulting lesser authentication performance causes an unacceptable amount of false acceptances and rejections, and prevents the use of proposed BCS and CB approaches in many domains [5].

Authorization, i.e., access control, is normally carried out after authentication in order to grant a user data and operating privileges. Many access control models have been proposed and widely deployed in order to grant users privileges in a systematic manner [9]. Discretionary Access Control (DAC) models [17], [18], where each user is given an explicit set of privileges, were once popular in commercial domains because of their flexibility and fine-granularity. Unfortunately, DAC models do not scale well as large numbers of users, each with their own set of privileges, become increasingly difficult to manage. Mandatory Access Control (MAC) models [19] introduced privilege levels. In a MAC model, a user is assigned a privilege level and granted access to all the data objects of equal or lower privilege level. Unfortunately, MAC models are not well-fit for high security domains where many data objects may be accessible by only a small set of corresponding users. Role-based Access Control (RBAC) models [10] have been widely accepted as the restrictions of DAC and MAC schemes were recognized. In RBAC models, privileges are assigned to groups of users based on their role within an organization. RBAC simplifies privilege management when a user's role in the organization changes and also facilitates complex privilege hierarchies with fine-granularity. As a result, RBAC models are well-suited for computing systems which group their users into hierarchical roles.

Few proposed systems have attempted to integrate authentication and authorization. The Cisco RADIUS protocol (The Remote Authentication Dial-In User Service) claimed to couple authentication and authorization [20]. Unfortunately, the protocol's authorization process, which is solely an access request to a server along with the username and password, is vague and weak.

III. BioCapsule Scheme Overview

In order to perform the novel, integrated AuthN-AuthZ operation, our proposed system leverages the provably secure and fusion-based Biometric-Capsule, or BioCapsule, (BC) scheme [7], [8]. As the BC scheme is central in performing the proposed, integrated AuthN-AuthZ operation, we provide a brief overview of the BC scheme. For additional details, please see [7], [8].

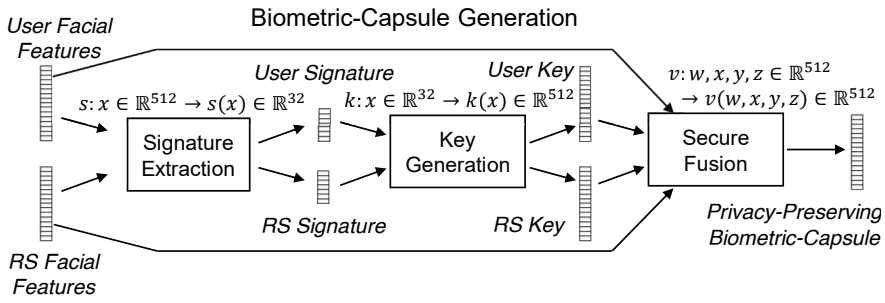

Fig. 1. Biometric-Capsule (BC) generation involving: signature extraction, key generation and secure fusion. Notice that the input of BC generation is a pair of biometric feature vectors (one belonging to a user and the other belonging to the user's corresponding RS). As no restrictions are made upon the biometric sampling, preprocessing, feature extraction or classification steps of the overall system, the BC scheme can be embedded into existing biometric systems.

A. Reference Subjects

The BC scheme involves the introduction of a Reference Subject (RS) [7]. An RS can be logical, such as an image, or a physical item, such as a doll. During system enrollment each user is assigned a corresponding RS. The biometrics of a user are then securely fused with the biometrics of their corresponding RS in order to generate a resulting BC. The resulting BC, rather than the user or RS biometrics, are stored by the authentication system. Later, during any authentication attempt, a user must present their own biometrics to the authentication system as well as their assigned RS. The system can again securely fuse the sampled biometrics of the user and the RS to form a new, query BC. This query BC is compared with the BC(s) stored by the system during enrollment in order to make an authentication decision. If an attacker steals a user's BC, the victim user can then revoke (or cancel) their stolen BC and re-enroll using a different RS for future authentication attempts [7].

As RSs are required for any authentication attempt, how a user provides their RS to the system becomes an important issue. In high-security scenarios, such as in military domains, the authentication system may require a physical RS to be carried by users in order to be authenticated. If logical (image) RSs are used, it would be possible for the system to store all RSs assigned to users and allow a query user to pick their RS from a list of all RSs at authentication time. Though these RS setups offer additional security by creating a Two-Factor Authentication (the user is responsible for providing their own biometrics as well as providing/selecting their RS), they lessen the usability benefits of biometrics-based authentication.

Therefore, a third setup is possible. If a user is required to simply input an assigned username (determined during enrollment) at authentication time, the BC scheme can be made fully transparent to users. The system can store all RSs and a username-to-RS mapping. Then, based on the username entered at authentication time, the appropriate RS can automatically be retrieved by the system and used for BC fusion. We leverage this user-friendly RS setup within our proposed AuthN-AuthZ system.

B. BioCapsule Generation

Regardless of what type of RS and method of maintaining RSs, BC generation always works the same. After biometric sampling, preprocessing and feature extraction are performed, the fusion-based BC generation is carried out. BC generation involves three steps (see Figure 1): signature extraction, key generation and secure fusion. Each of these steps involve one-way functions which result in the overall BC generation being one-way. Furthermore, each of these steps serve to successfully mask the contributions of a user's and RS's biometric features toward a resulting BC [7], [8].

First, signature extraction is carried out. This task is done using the three-level averaging method [21]. The method first reshapes an input biometric feature vector into a matrix. For instance, a $\mathbb{R}^{512}$ biometric feature vector (the same shape as the feature vectors used in our proposed system) is reshaped into a $\mathbb{R}^{32x16}$ feature matrix. Second, two averaging convolutions using kernels of different sizes are applied to the reshaped matrix. In our proposed system, we applied kernels of sizes $\mathbb{R}^{3x3}$ and $\mathbb{R}^{5x5}$. The absolute difference between these two convolutions is then computed. Finally, a row-wise average is applied to the absolute difference. In the end, the three-level averaging of the $\mathbb{R}^{512}$ vector results in a $\mathbb{R}^{32}$ signature (i.e. $s : x \in \mathbb{R}^{512} \rightarrow s(x) \in \mathbb{R}^{32}$). This process represents a one-way function as, given a set of values, it is easy to obtain an average. However, given an average, infinitely many sets of corresponding values can be derived.

Next, key generation is carried out. This is accomplished by scaling each signature value by a power of 10 (e.g., 10^2 in our system) and then rounding the resulting values to integers. Each of the resulting 32 scaled and rounded signature values are then used as seeds for a random number generator (RNG). Each seed is used in order to produce 16 uniformly random values of $\{-1, 1\}$. These values are combined into a single $\{-1, 1\}$ key vector of size $\mathbb{R}^{512}$ (the same size as the initial feature vector). Therefore, key generation maps a signature vector $\mathbb{R}^{32}$ of size to a key vector of size $\mathbb{R}^{512}$ (i.e $k : x \in \mathbb{R}^{32} \rightarrow k(x) \in \mathbb{R}^{512}$). It can be seen that this process is also one-way as, given a set random of uniformly random values $\{-1, 1\}$, one cannot derive the seed which was initially given to the RNG.

After extracting signatures and keys from the user's feature

vector and the feature vector of the user's RS, BC secure fusion is performed. BC secure fusion can be modeled using the following equation:

$$F^{User,RS} = F^{User} * K^{RS} + F^{RS} * K^{User} \quad (1)$$

where F^{User} and F^{RS} are the user and RS features respectively, K^{User} and K^{RS} are the user and RS keys respectively, $*$ is an element-wise multiplication, $+$ is a simple vector addition and $F^{User,RS}$ is the resulting BC [8]. It should be noted that the resulting BC is of shape $\mathbb{R}^{512}$ (the same shape as the original input user and RS feature embeddings). Therefore, secure fusion maps four vectors (user feature, RS feature, user key and RS key) of size $\mathbb{R}^{512}$ to a single BC of size $\mathbb{R}^{512}$ (i.e. $v : w, x, y, z \in \mathbb{R}^{512} \rightarrow v(w, x, y, z) \in \mathbb{R}^{512}$).

A few aspects of the BC generation process should be noted. First, no biometric feature information is lost during BC generation. As keys only contain values $\{-1, 1\}$, the element-wise products in BC fusion only negate or do not effect the values within the feature embeddings. Second, the user and RS biometric features contribute equally to a resulting BC, (i.e. no additional weight is given to either the user or RS contributions within Equation 1). Third, the BC scheme is not specific to any biometric modality. The BC scheme is a general approach which can be applied to any biometric modality. In this work we secure a facial authentication system with the BC scheme, but the BC scheme could easily be extended to a different biometric modality. In fact, extensive experiments have shown that the BC scheme can be effectively used to secure both iris and facial authentication systems [7], [8].

One other notable aspect of BC generation is its modularity and flexibility. As illustrated in Figure 1 and demonstrated by Equation 1, the BC scheme requires no fixed biometric sampling, preprocessing, feature extraction, or classification techniques in order to accommodate it. The BC scheme simply requires the introduction of an RS and the feature embeddings of both a user and their corresponding RS. As the input feature embeddings would be used as biometric templates within an unsecured system, the BC scheme can be embedded into existing, underlying biometric authentication systems. This gives the BC scheme the major advantage that it can be be used alongside the most recent, state-of-the-art deep learning-based techniques [8].

C. BioCaspule Security Analysis

Generated BCs are provably secure and privacy-preserving against several types of attacks. Here we provide a brief overview of the BC scheme's resistance to the most common types of biometric template attacks that could arise (for the detailed security proofs please see [7]). It should be noted that the BC scheme is a biometric template security scheme and, therefore, it does not consider attacks against a biometric authentication unrelated to biometric template security (e.g. biometric spoofing attacks, man-in-the-middle attacks, deep learning-based adversarial example attacks, etc.). Instead, the BC scheme is used to robustly secure user biometric templates such that they do not reveal the user's sensitive, personal information (ethnicity, gender, age, heath condition, etc.) upon analysis.

Here, we consider several cases in which an attacker steals a victim user's BC and then tries to derive the user's biometric feature embedding. Using this feature embedding, the attacker can easily reveal the user's sensitive, personal information.

The first case which we consider is the case in which an attacker steals only the victim user's BC, $F^{User,RS}$, and attempts to derive the user's feature embedding, F^{User}. As shown in Equation 1, the attacker will be presented with an underdetermined equation. Therefore, it is impossible for the attacker to derive F^{User}.

The second case which we consider is the case in which the attacker steals both the victim user's BC and corresponding RS. As the attacker has the victim user's RS, they can derive both F^{RS} and K^{RS}. Then, by rearranging Equation 1, the attacker is forced to guess each key value within K^{User}. As each key value is $\{-1, 1\}$, this results in two possible feature values for each value within F^{User}. Since K^{User} and F^{User} each contain 512 values, the attacker will be able to derive $O(2^{512}) \approx O(10^{154})$ possible victim user feature vectors. To derive all such possible feature vectors and use them to determine the user's sensitive, personal information is computationally infeasible [7].

The third case which we consider is the case in which the attacker steals multiple BCs of a victim user (or multiple victim users). Similar to the first case, this will present the attacker with an underdetermined system of equations. Therefore, it is impossible for the attacker to derive F^{User}.

The fourth and final case which we consider is the case in which the attacker steals multiple BCs of a victim user (or multiple victim users) and the victim user's (or victim users' shared) corresponding RS. This will result in a sub-cases of the second case for each stolen BC. As noted previously, this presents the attacker with a computationally infeasible amount of possible feature vectors to derive for each stolen BC.

From these cases it can be seen that multiple users may be assigned the same shared RS with no loss in the robust security and privacy-preserving benefits of the BC scheme. These security and privacy-preserving benefits offered by the BC scheme directly address user concerns. Users can take advantage of the usability benefits provided by the BC-embedded biometric authentication and, at the same time, feel assured their biometric information is robustly secured from any attacker.

IV. Reference Subject/Role-based Access Control Model

In addition to the BC scheme, the AuthN-AuthZ system will leverage a hierarchical Reference Subject/Role-based Access Control (RS-RBAC) model (e.g., Figure 2). Such a model can be represented as an directed acyclic graph (DAG).

Vertices within the DAG represent roles within a computing system with corresponding data and operating privileges. Each role is assigned a corresponding RS image to be used during BC-embedded facial authentication. Each user is assigned a

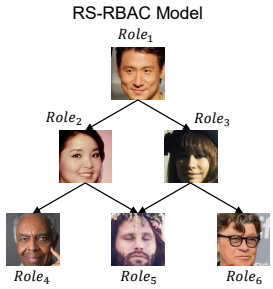

Fig. 2. Example hierarchical RS-RBAC model used in our experiments. Each role's corresponding RS image is shown.

role and the role's corresponding RS based on their activities within the computing system and the set of privileges that the user should be granted.

Directed edges within the DAG denote hierarchical relations between roles. If a path exists from a source role (ancestor) to a destination role (descendant), the source role will be granted all its own corresponding privileges as well as the privileges of the descendant role. Thus, ancestor roles gain all the data and operating privileges of all their descendants.

The resulting RS-RBAC model enables an administrator to exercise fine-grained access controls and model complex relationships between users. Such an administrator will be responsible for maintaining the RS-RBAC. This includes maintaining: the privileges of each role within the RS-RBAC, the user-role assignments (denoted by the users' username within the system), the RS images assigned to roles and the hierarchical relationships between roles. All these tasks can be carried out easily and efficiently using an RS-RBAC implemented through the use of a key management technique such as [22].

V. AuthN-AuthZ System Design

Utilizing the BC scheme and the RS-RBAC model, we propose a highly user-friendly, secure and privacy-preserving Biometric-Capsule-embedded facial authentication system which is capable of performing an integrated AuthN-AuthZ operation. The overall enrollment and authentication workflows of the proposed AuthN-AuthZ system are shown in Figure 3. As shown in Figure 3, the facial enrollment and authentication workflows are comprised of multiple steps including: biometric sampling, RS retrieval from the RS-RBAC, biometric preprocessing (including facial detection, alignment and segmentation), feature extraction, feature representation, BC generation and finally classification. Below, we discuss each of these steps in detail.

A. Username, Biometric Sampling and RS-RBAC

The BC-embedded facial authentication workflow begins by prompting a user for a username and sampling the user's facial biometrics.

Before enrollment can be performed, the system administrator responsible for maintaining the RS-RBAC must create a username for a new user and assign them to a role within the RS-RBAC (or create a new role for the user in the RS-RBAC). The system administrator will then need to communicate the username with the new user. Later, during enrollment or authentication, the new user will enter the username assigned to them by the system administrator. The entered username will be used to retrieve the user's RS image from the RS-RBAC through the use of a username-to-role mapping. As a user only needs to remember their username and their RS is automatically retrieved from the RS-RBAC, BC generation is made completely transparent to users.

To sample a user's facial biometrics, we utilize the embedded, front-facing camera of a smart device or a webcam connected to a laptop/desktop machine. The system captures a video stream from these biometric sensors. Each frame of the captured video stream can then be leveraged by the subsequent steps within the facial authentication workflow.

One notable advantage of using the frames of a video stream is that it enables rapid system enrollment and authentication. During enrollment, several images can be rapidly captured by the biometric sensor, forwarded to subsequent steps of the facial authentication workflow and finally transformed into BCs. The system then stores only the transformed BCs for privacy-preservation. During authentication, the system needs to acquire a single facial image, which will be processed, fused into a BC and compared with stored BCs for authentication. Due to the unconstrained nature of how many users configure their webcam or smart device's embedded camera, many frames of the video stream will be deemed unsuitable for authentication (as discussed later in this section). Since the video stream captures many frames each second, unsuitable frames can quickly be discarded and replaced by suitable frames.

In the case that a user enters a username which has not been created by the system administrator, the user's biometrics are still sampled, but the user will be notified that the enrollment or authentication attempt was unsuccessful. We do not explicitly inform the user that the entered username was invalid as it may cause information leakage about usernames registered in the system.

B. Preprocessing

After a biometric sample is captured by the system and an RS image is retrieved, biometric preprocessing will take place. In our system, the state-of-the-art, widely accepted and deep learning-based Multi-Task (Cascaded) Convolutional Neural Network (MTCNN) technique [25] is utilized for preprocessing. This technique uses a cascade (ensemble) of three Convolutional Neural Networks (CNNs). The first CNN in the cascade is the Proposal Network (P-Net), which will find potential regions of interest in an image which may contain a face. The second CNN is the Refinement Network (R-Net), which will reject regions proposed by the P-Net which do not include a face. The third CNN is the Output Network (O-Net), which will further reject falsely proposed facial regions given by the R-Net. In addition, for each region which it accepts as correct facial detections, the O-Net will identify the positions of five facial landmarks. Therefore, the output(s)

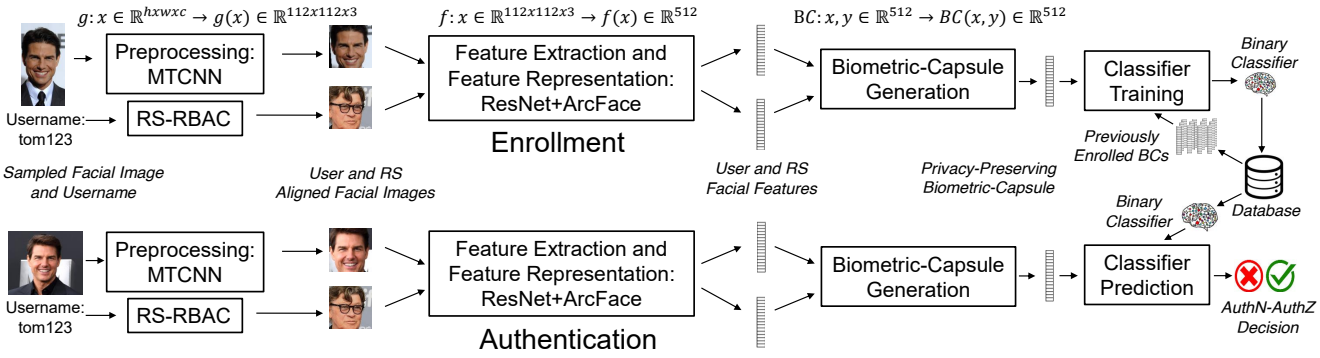

Fig. 3. Overall proposed AuthN-AuthZ system (enrollment work-flow shown on top and authentication work-flow on bottom): A user begins enrollment or authentication by providing their username and allowing the system to sample their facial biometric traits. The RS image corresponding to the given username's role is then retrieved from the RS-RBAC. Meanwhile, the sampled facial image is preprocessed. User and RS facial features are then extracted and embedded using a ResNet model [23] and the ArcFace method [24], respectively. Next, the user and RS features are fused in order to generate a privacy-preserving BC. Finally, the BC is used in enrollment to train a binary classifier or classified during authentication in order to make an authentication decision.

of the MTCNN cascade are facial detection regions, each with the positions of five corresponding facial landmarks. The five facial landmarks represent the location of the right eye, the left eye, the tip of the nose, the left corner of the mouth, and the right corner of the mouth. Our system utilizes an open-source implementation of MTCNN provided by [26].

If the MTCNN detects no faces within an image, the image is merely considered unsuitable for authentication and is rejected without further consideration. If such a case arises, another biometric signal is captured from the video stream and facial preprocessing is restarted. If one or more faces are detected, the center-most face within an image is then considered further, while other faces are discarded. The detected face (or center-most detected face) is then aligned. Alignment is performed using an Affine transformation to rotate the image such that the angle between the detected right eye and left eye facial landmarks is reduced to zero. After alignment, segmentation is performed upon the aligned image. This involves cropping the resulting facial region from the rotated image. We crop the facial region along with a 42-pixel margin, which is chosen to ensure the preservation of useful biometric information such as hairline, jawline, facial hair, etc. and reshape the cropped facial image to shape $\mathbb{R}^{112x112x3}$. As a result, the system's preprocessing steps map an arbitrary size input image to a fixed sized aligned image (i.e. $g : x_i \in \mathbb{R}^{hxwx3} \rightarrow g(x_i) \in \mathbb{R}^{112x112x3}$).

It should be noted that, in high security scenarios, it may be appropriate for system designers to choose to reject any image containing multiple faces. We choose to use the center-most face due to the nature of the experimental datasets used in Section VI.

C. Feature Extraction and Feature Representation

Following facial preprocessing, facial feature extraction and representation steps are preformed. To perform feature extraction, we use a 100-layer, extremely deep Residual Neural Network (ResNet) [23] architecture CNN. The ResNet architecture is a widely accepted and state-of-the art CNN architecture. ResNets feature "skip-connections" which combine the output of convolutional layers with their original inputs. As a result, these skip-connections enable ResNets to learn convolutional layers which can act as identity functions. Furthermore, the skip-connections allow ResNets to be made deeper by addressing issues inherent in deep architectures, such as the exploding/vanishing gradient problem [23]. The ResNet model serves to map aligned input images of size $\mathbb{R}^{112x112x3}$ to feature vectors of size $\mathbb{R}^{512}$ (i.e. $f : x_i \in \mathbb{R}^{112x112x3} \rightarrow f(x_i) \in \mathbb{R}^{512}$).

Unfortunately, the Cross-Entropy Loss typically used to train CNN models does not produce discriminative feature vectors (i.e., given two biometric feature vectors, it is easily discernible if the they belong to the same subject or not). Cross-Entropy Loss only aims to separate the feature vectors of subjects known at training time such that the feature can be partitioned using subject-based decision boundaries. The resulting feature vectors are often not sufficiently discriminative for facial recognition tasks (i.e. it may be ambiguous if two feature vectors belong to the same subject) and it does not enable the CNN to produce feature vectors that generalize well to subjects unseen at model training time.

In order to ensure features yielded by the ResNet model are discriminative, the ResNet model used in our system is trained using the recently proposed and state-of-the-art ArcFace Loss [24]. Through modification of standard Cross-Entropy Loss, ArcFace Loss fine-tunes a CNN during training in order to maximize the discriminative properties of extracted feature vectors. ArcFace Loss [24] modifies Cross-Entropy Loss in order to pull each subject's feature vectors toward a respective class-center. This provides intra-class compactness. Furthermore, it introduces a margin penalty which serves to create a linear margin between each subject's feature vector cluster. This, on the other hand, provides inter-class variability. As a result, ArcFace creates a feature representation in which distance between feature vectors directly corresponds to facial similarity. The ArcFace technique has been shown to yield extremely discriminative facial feature vectors. In fact, at the time of writing, ArcFace has achieved the top results on the challenging Labeled Faces in the Wild (LFW) [27] face

verification benchmark.

In our system, feature extraction and representation works by simply passing an input preprocessed $\mathbb{R}^{112x112x3}$ facial image through an ArcFace Loss trained ResNet model and receiving an output feature embedding of size $\mathbb{R}^{512}$ (i.e. $f : x_i \in \mathbb{R}^{112x112x3} \rightarrow f(x_i) \in \mathbb{R}^{512}$). We leverage an open-source, pre-trained model given by the ArcFace authors [26].

D. BioCapsule Generation

After extracting feature embeddings from both the user's biometric sample and the user's RS image, the two feature embeddings are ready to be securely fused into a resulting, privacy-preserving BC. As explained in Section III this requires the system to carry out three steps: signature extraction, key generation and secure fusion. Please see Section III-B for the explicit details of how these steps are carried out within our proposed system. As a result of BC generation, the input $\mathbb{R}^{512}$ user and RS feature embeddings are fused into a corresponding $\mathbb{R}^{512}$ BC (i.e. $BC : x, y \in \mathbb{R}^{512} \rightarrow BC(x, y) \in \mathbb{R}^{512}$

E. Classification

A resulting BC can then be used for system enrollment or authentication. During enrollment, generated BCs are simply stored by the system for use in later authentication decisions. During authentication, a generated BC must be compared to the enrolled BCs of the subject whom a query user claims to be. This comparison, which represents a binary classification problem, will result in an authentication decision.

For this reason, in addition to storage of each user's corresponding BCs, a binary classifier must be trained and stored for each subject. These binary classifiers are trained during enrollment and later used anytime a query user attempts to be authenticated as the binary classifier's corresponding enrolled subject. Given a query BC, the binary classifier will output a predicted probability that the BC does indeed belong to the enrolled subject and, furthermore, that the correct RS (which the enrolled subject was assigned) was used in the query BC's generation.

More specifically, binary classifier training begins during enrollment after the system has generated several BCs for a new subject to be enrolled. Using all previously enrolled BCs and this new set of BCs, a Logistic Regression (LR) classifier is trained for the new subject. The LR takes as input a $\mathbb{R}^{512}$ vector (the same size as the generated BCs) and outputs a scalar value $[0, 1]$. After training the LR classifier, this scalar output denotes the predicted probability that a query BC should be successfully authenticated as the newly enrolled subject. As previously noted, since any query BC securely encapsulates both the feature embedding of a query user and a corresponding RS, a query BC should only be successfully authenticated when both the query user is indeed the enrolled subject they claim to be, and when the query user used the enrolled subject's corresponding RS in BC fusion. If the encapsulated feature embedding of the query user is not sufficiently similar to the enrolled subject's feature embeddings or if the wrong RS is used, the classifier should reject the query BC.

VI. Experiment

In this section, we perform extensive experiments to evaluate the performance of the proposed BC-embedded facial authentication system and AuthN-AuthZ operation. The experiments consider authentication performance under adversarial test cases and demonstrate the proposed system's robust performance and flexibility. Furthermore, we compare the performance of the proposed BC-embedded system to the underlying version of the authentication system which does not perform BC fusion (and instead performs authentication using unsecured feature vectors). This comparison allows us to directly characterize the BC scheme's effects on underlying authentication performance.

A. Dataset

We utilized two benchmark facial biometrics datasets to demonstrate the performance of the proposed system. The first dataset is the constrained Georgia Tech Face Database (GTDB) [28]. The dataset features variation in facial pose, facial expression and facial accessories. The dataset is quite constrained as all photos contain a single subject within a well-illuminated office environment.

The second experimental dataset is a subset of the popular benchmark Labeled Faces in the Wild (LFW) dataset [27]. Unlike many constrained facial biometric datasets, the highly unconstrained LFW dataset features variation in facial pose, facial expression, facial accessories, illumination, setting, number of faces within each image and facial occlusions.

We began the experiment by filtering the datasets. Any subject from the datasets that do not have at least five images are removed from the experimental datasets. This filtering process ensures that each subject will have at least four images for training (enrollment) and at least one image for testing (authentication). In the end, the filtered LFW dataset is reduced to a subset of its original subjects, and, as a result, it contains 423 subjects and 5,985 images. No subjects are filtered out of the GTFB which contains 50 subjects and 750 images.

For each dataset we perform stratified five-fold cross validation. This involves splitting the dataset into five random folds. Each subject's images are equally distributed across each of the five folds. Then, a single fold is selected as the testing set while the remaining four folds are combined and used as the training set. This process is repeated five times such that each fold is used as the testing set once.

B. RS Assignment using the RS-RBAC Model

To test the proposed AuthN-AuthZ system's ability to support different RS-RBAC models, we performed two test cases. In these two cases, users are assigned to roles (and their corresponding RSs) using the example RS-RBAC model shown in Figure 2. In each case, we used different role assignment probability distributions.

TABLE I
RS-RBAC ROLE PROBABILITIES AND SUBJECT ASSIGNMENTS

Dataset	Role Distribution	$Role_1$		$Role_2$		$Role_3$		$Role_4$		$Role_5$		$Role_6$	
		Prob.	Subj.	Prob.	Subj.	Prob.	Subj.	Prob.	Subj.	Prob.	Subj.	Prob.	Subj.
GTDB	Balanced	16.67%	6	16.67%	6	16.67%	10	16.67%	12	16.67%	9	16.67%	7
	Unbalanced	5%	3	10%	5	15%	11	20%	9	25%	13	25%	9
LFW	Balanced	16.67%	80	16.67%	63	16.67%	71	16.67%	81	16.67%	57	16.67%	71
	Unbalanced	5%	27	10%	43	15%	60	20%	73	25%	111	25%	109

TABLE II
AUTHENTICATION RESULTS

Dataset	Template Type	Role Distribution	FP	FN	ACC	FAR	FRR
GTDB	ArcFace Feature	N/A	0	0	100%	0%	0%
	Biometric-Capsule	Balanced	0	0	100%	0%	0%
		Unbalanced	0	0	100%	0%	0%
LFW	ArcFace Feature	N/A	5	5	99.9996%	0.0002%	0.0835%
	Biometric-Capsule	Balanced	31	29	99.9976%	0.0012%	0.4845%
		Unbalanced	38	50	99.9965%	0.0015%	0.8354%

The first role assignment distribution gives unbalanced likelihood that a user will be assigned to each role. This test is quite meaningful as ancestor roles within the RS-RBAC hierarchy will have access to both their corresponding data and operation privileges as well as their descendants' privileges. It is therefore reasonable to assume that, within many applications, less users will be assigned these more privileged roles. To reflect this, we defined a probability that each user will be assigned to each of the roles (and it's corresponding RS) within the hierarchy. We then randomly assigned each of the two datasets' subjects roles and corresponding RSs using these probabilities.

We also tested the BC system using a balanced role assignment distribution where user assignment to each role is equally likely. The probability of each role assignment and the resulting number of subjects assigned to each role is shown in Table I.

C. Enrollment

We began enrollment by performing preprocessing upon each image within the training set as described in Section V-B. After preprocessing, we performed data augmentation. This involves flipping each preprocessed training image across its vertical axis. We included each flipped image in the training set, and, as a result, doubled the size of the training set. We then performed feature extraction and representation upon each of the preprocessed images within the augmented training set as described in Section V-C. The resulting features were then used for BC generation as described in Section III-B.

A set of training BCs were formed by fusing each training feature with each RS within the RS-RBAC. Next, we partitioned the resulting BC training set into positive and negative training sets with respect to each subject. Each subject's positive training set is made up of all BCs formed by fusing the subject's training features with the feature of the subject's corresponding RS (subject-match/RS-match). A subject's negative training set includes all BCs formed by fusing the features of any other subjects with the feature of the subject's corresponding RS (subject-mismatch/RS-match). Since a user's RS is automatically selected based on their username, subject-match/RS-mismatch and subject-mismatch/RS-mismatch BCs are not possible.

A set of testing BCs were formed by fusing each testing feature with each RS within the RS-RBAC. We assume the adversarial case where each user's username is public and known to attackers. Therefore, each subject's testing set contained all testing BCs formed using the subject's correct RS during fusion.

For each subject, we then train an LR binary classifier (as described in Section V-E) using the subject's respective positive and negative training sets. Since the positive training set of each subject is much smaller than the negative training set, it is necessary to apply additional weight to misclassifications of positive examples during the training process. To accomplish this, we weigh positive examples inversely proportional to their frequency within the entire BC training set. We also do the same for the negative set to lower their weight. Finally, the resulting trained LR models are used to predict each their subjects' corresponding testing sets.

D. Integrated Authentication and Authorization

The results of the authentication experiment are displayed in Table II. We report the five-fold cross-validation macro-average of authentication accuracy (ACC), false acceptance rate (FAR) and false rejection rate (FRR). We also report the total amount of false positives (acceptances) (FP) and false negatives (rejections) (FN).

In the case of the constrained GTDB dataset, 37,500 authentication attempts are made across the 5-fold cross-validation experiment. These 37,500 authentication attempts are comprised of 750 genuine authentication attempts which the system should accept and 36,750 authentication attempts which the system should reject. The underlying system and BC-embedded system (with both balanced/unbalanced role distributions) are able to perfectly perform classification across the five-fold cross-validation. This perfect performance would be expected from the deep learning-based underlying system (which is unable to perform the AuthN-AuthZ operation). It is quite notable that the BC-embedded system (under each of the RS/role distributions) is able match this perfect performance. This demonstrates that the BC scheme is able to successfully leverage the discriminative and state-of-the-art robustness of the underlying deep learning-based system while at the same

time securing user privacy and facilitating the AuthN-AuthZ operation.

For the unconstrained LFW dataset, 2,531,655 total authentication attempts are made. These authentication attempts are comprised of 5,985 genuine authentication attempts which the system should accept and 2,525,670 authentication attempts which the system should reject. The underlying system is able to slightly outperform the BC-embedded AuthN-AuthZ system. This slight decrease in performance is as expected [8], but the BC-embedded system has great advantage over the underlying system in terms of privacy-preservation and the ability to facilitate the AuthN-AuthZ operation. The underlying system allows 5 FP and 5 FN. This yields an FAR of 0.0002% and FRR of 0.0835%. The BC system allows 31 FP and makes 25 FN in the case of balanced role distribution. This yields a FAR of 0.0012% and a FRR of 0.4845%. In the case of unbalanced role distribution, the BC system makes additional misclassifications, i.e. 38 FP and 50 FN. These additional misclassifications yield a FAR of 0.0015% and a FRR of 0.8354%.

The difference in performance between the two BC systems can be explained by their corresponding role distributions. In the case of the unbalanced role distributions, more subjects are assigned to a few roles (rather than the balanced distribution where each role is assigned an approximately equal amount of subjects). In paper [8], authors suggest that BC-embedded system performance will diminish as larger groups of users share common RSs. Since RSs and user features contribute equally to BCs, BCs generated using the same RS may have lesser inter-class distinguishability despite being generated using different users' feature embeddings.

Although the BC-embedded AuthN-AuthZ performance does not match that of the underlying system, its comparable performance is quite notable. By raising FAR by 0.0013% in the worse case (i.e. LFW underlying system compared with the BC-embedded system with unbalanced role distribution), the AuthN-AuthZ is able to protect the privacy of user's biometric attributes. Though the FRR is increased by 0.7159% in the worst case, false rejections are not as great a concern as false acceptances. If a user is falsely rejected by the system, re-authentication attempts can be made quickly via the efficient biometric sampling described in Section V-A. It should be noted that, from a security point-of-view, misclassifications in the AuthN-AuthZ system are equivalent to misclassifications in a system incapable of the AuthN-AuthZ operation. Since RSs are automatically selected by the system, an attacker who causes a misclassification is only granted the access rights of the victim user they are authenticated/authorized as. This would be the same in a system which does not perform the AuthN-AuthZ operation.

VII. Conclusion

In this work we have proposed a highly user-friendly, privacy-preserving, biometrics-based authentication system. This system addresses many traditional usability issues posed by knowledge and object-based authentication methods. The system makes use of the provably secure and privacy-preserving BC scheme in order to address the privacy concerns associated with the use of biometric data which have recently had negative impacts upon users' perceptions and acceptance of biometrics-based technologies.

Our proposed system is able to facilitate authorization using a RS-RBAC model, and, furthermore, is able to perform a novel, integrated AuthN-AuthZ operation. In our experiment we have shown that the proposed BC-embedded facial authentication system and AuthN-AuthZ operation has comparable performance to an underlying unsecured system. In fact, from an authorization point of view, the security of the AuthN-AuthZ operation is equivalent to a system which uses traditional, independent authentication and authorization processes. Furthermore, the BC generation and AuthN-AuthZ operation can be carried out in a manner fully transparent to any user. This result is quite encouraging and provides strong motivation for the use of the novel, integrated AuthN-AuthZ operation.

Acknowledgments

This work is partially supported by a U.S. National Science Foundation grants (NSF OAC-1839746 & DGE-2011117) and also the NSF Jetstream [29]/XSEDE [30] project (ACI-1445604 & OCI-1053575). We thank Mr. Jeremy Fischer and Mr. David Hancock for their assistance with allocating needed computing resources and porting the developed VM image on JetStream, which was made possible through the XSEDE Extended Collaborative Support Service (ECSS) program.

References

[1] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Trans. on CSVT*, vol. 14, no. 1, pp. 4–20, Jan 2004.
[2] J. Chamary, "No, apple's face id is not a 'secure password'," *www.forbes.com/sites/jvchamary/2017/09/18/security-apple-face-id-iphone-x/#4adbcafd4c83*, September, 2017 Accessed in Sep. 2019.
[3] P. Mozur, "Inside china's dystopian dreams: A.i., shame and lots of cameras," *https://www.nytimes.com/2018/07/08/business/china-surveillance-technology.html*, July 8, 2018. Accessed in Sep. 2019.
[4] S. Musil, "Microsoft calls for regulation of facial-recognition technology," *https://www.cisco.com/c/en/us/support/docs/security-vpn/remote-authentication-dial-user-service-radius/13838-10.html#comp_auth*, May 2, 2016. Accessed in February 2019.
[5] C. Rathgeb and A. Uhl, "A survey on biometric cryptosystems and cancelable biometrics." *EURASIP J. Info. Security*, vol. 2011, p. 3, 2011.
[6] U. Uludag, S. Pankanti, and et. al., "Biometric cryptosystems: issues and challenges," *Proc. of the IEEE*, vol. 92, no. 6, pp. 948–960, 2004.
[7] Y. Sui, X. Zou, E. Y. Du, and F. Li, "Design and analysis of a highly user-friendly, secure, privacy-preserving, and revocable authentication method," *IEEE Trans. on Computers*, vol. 63, no. 4, pp. 902–916, 2014.
[8] T. Phillips, X. Zou, F. Li, and N. Li, "Enhancing biometric-capsule-based authentication and facial recognition via deep learning," in *Proc. of ACM SACMAT*, ser. SACMAT '19. New York, NY, USA: ACM, 2019, pp. 141–146.
[9] R. S. Sandhu and P. Samarati, "Access control: principle and practice," *IEEE communications magazine*, vol. 32, no. 9, pp. 40–48, 1994.
[10] R. S. Sandhu, E. J. Coyne, H. L. Feinstein, and C. E. Youman, "Role-based access control models," *Computer*, vol. 29, no. 2, pp. 38–47, 1996.
[11] S. Sinclair and S. W. Smith, "What's wrong with access control in the real world?" *IEEE Security Privacy*, vol. 8, no. 4, pp. 74–77, July 2010.
[12] T. Phillips, "Biocapsule github repository," *https://github.com/phillity/BioCapsule*.
[13] E. Eidinger, R. Enbar, and T. Hassner, "Age and gender estimation of unfiltered faces," *IEEE TIFS*, vol. 9, no. 12, pp. 2170–2179, Dec 2014.
[14] X. Boyen, "Reusable cryptographic fuzzy extractors," in *Proc. of ACM CCS*, ser. CCS '04. New York, NY, USA: ACM, 2004, pp. 82–91.

[15] H. Lu, K. Martin, F. Bui, K. N. Plataniotis, and D. Hatzinakos, "Face recognition with biometric encryption for privacy-enhancing self-exclusion," in *16th Inter. Conf. on Dig. Sig. Processing*, July 2009, pp. 1–8.
[16] M. Savvides, B. Kumar, and et. al., "Cancelable biometric filters for face recognition," in *Proc. of ICPR*, vol. 3, Aug 2004, pp. 922–925.
[17] M. A. Harrison, W. L. Ruzzo, and J. D. Ullman, "Protection in operating systems," *Comm. of the ACM*, vol. 19, no. 8, pp. 461–471, 1976.
[18] D. D. Downs, J. R. Rub, K. C. Kung, and C. S. Jordan, "Issues in discretionary access control," in *IEEE S&P*. IEEE, 1985, pp. 208–208.
[19] D. E. Denning, "A lattice model of secure information flow," *Communications of the ACM*, vol. 19, no. 5, pp. 236–243, 1976.
[20] Cisco, "TACACS+ and RADIUS comparison," *https://www.cisco.com /c/en/us/support/docs/security-vpn/remote-authentication-dial-user-service-radius/13838-10.html#comp_auth*, 2008/1/4 Accessed in Sep. 2019.
[21] Y. Du, R. Ives, and et. al., "Use of one-dimensional iris signatures to rank iris pattern similarities," *Opt. Eng.*, vol. 45, no. 3, Mar. 2006.
[22] M. Atallah, M. Blanton, and et. al., "Dynamic and efficient key management for access hierarchies," *ACM Trans. ISS*, vol. 12, no. 3, pp. 18:1–43, Jan 2009.
[23] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *2016 IEEE CVPR*, pp. 770–778, 2015.
[24] J. Deng, J. Guo, and S. Zafeiriou, "Arcface: Additive angular margin loss for deep face recognition," *ArXiv*, vol. abs/1801.07698, 2018.
[25] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1499–1503, Oct 2016.
[26] D. Insight. (2019) Face analysis project using mxnet repository. Available at: https://github.com/deepinsight/insightface.
[27] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," no. 07-49, October 2007.
[28] M. H. H. Ara V. Nefian, Mehdi Khosravi, "Real-time detection of human faces in uncontrolled environments," pp. 3024 – 3024 – 9, 1997.
[29] C. Stewart, T. Cockerill, I. Foster, and D. H. et al., "Jetstream: a self-provisioned, scalable sci. and eng. cloud environment," *XSEDE'15 Conf.: Sci. Adv. Enabled by Enhanced Cyberinfra.*, pp. 1–8, 2015.
[30] J. Towns and T. C. et al., "XSEDE: Accelerating scientific discovery," *Computing in Science & Engineering*, vol. 16, no. 5, pp. 62–74.