IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-7, July 2019 E-ISSN: 2347-2693

Risk-Based Authentication using Autoencoders

Bharat Sharma[1*], Sidharth Singh[2]

[1]Department of Information Technology, AIT College, Pune, India
[2]Department of Science and Technology, MM University, Ambala, India

[*]*Corresponding Author: sidharthsingh1806@gmail.com, Tel.: +91-8251012333*

DOI: https://doi.org/10.26438/ijcse/v7i7.155160 | Available online at: www.ijcseonline.org

Accepted: 11/Jul/2019, Published: 31/Jul/2019

***Abstract*—** Verification gives a way to check the authenticity of a client attempting to get to any classified or delicate data. The requirement for ensuring secure information facilitated on the web has been rising exponentially as associations are moving their applications on the web. Static techniques for validation can't totally ensure the validity of a client. This has prompted the advancement of multifaceted validation frameworks. Risk-based validation; a type of multifaceted verification adjusts as per the risk profile of the clients. This paper advances the plan of risk motor incorporated with the framework to inspect the client's past login records and produce an appropriate example utilizing AI calculations to figure the risk dimension of the client. The risk level further chooses the confirmation technique that the client will be tested with. In this manner the versatile verification model aides in giving a more elevated amount of security to its clients.

***Keywords*—**User Metadata, Risk Metadata, Authentication System

I. INTRODUCTION

Shortcomings in secret phrase-based confirmation have been known for quite a while. They extend from frail and simple to figure passwords or secret key re-use to being helpless to phishing assaults. In any case, passwords are the dominating validation system sent by online administrations today. To expand the clients' security, administration administrators should actualize extra measures. Two-factor validation (2FA) is one broadly-offered measure that improves account security, however, is somewhat disagreeable (for example in January 2018, under 10 % of dynamic Google records utilized 2FA). Risk-based authentication (RBA) is a methodology that expands security with negligible effect on client association, and in this way can possibly give secure confirmation with great ease of use.

Client's accreditations utilized for validation are classified by what the client knows, what the client has, and what the client is. Passwords fall under the class of what the client knows, while tokens like dongle/mobile phone are instances of what the client has. Biometrics is a case of what the client is. Passwords can without much of a stretch be hacked and tokens can be stolen or reused. Biometric applications are costly to actualize. Be that as it may, a mix of these elements can advance higher security to online frameworks. Such a type of verification is called multifaceted confirmation.

Risk-based validation is one kind of multifaceted verification that adjusts as indicated by the client's risk profile. A risk profile is acquired by looking at the client's profile recovered at the season of login and past login records of the client. A model, portraying the client's conduct, is worked by a risk motor that is incorporated with the risk-based verification framework. The proposed work uses AI calculations to construct such a model. AI calculations are prepared to take in examples from accessible information and anticipate the obscure worth when furnished with a new arrangement of information.

Contingent upon the risk profile created, the client is tested with various verification techniques [1]. In this manner, a real client isn't required to pass different elements of confirmations to prove his genuineness, while a suspicious user needs to pass all the verification techniques he is tested with. This guarantees the framework is usable and security instead of most of the current frameworks that strike an exchange off between ease of use and security.

The rest of the paper is organized in the accompanying way segment 2 briefs about the current research works, breaking down their qualities and downsides. The detailed elaborations of the proposed technique are inspired in segment 3. Segment 4 introduces the test results and segment 5 exhibits the end.

II. RELATED WORK

A fundamental characteristic for rising danger based verification advancements is to utilize the client's conduct behavioural data alongside character and logical data in the risk of the boarding procedure. In any case, as far as anyone

is concerned, none of the works distributed so far in the exploration writing has considered the client personal conduct standards in the risk the executive's procedure. A large portion of the distributed papers on using behavioural pattern in risk-based authentication is industry whitepapers [2]. Similarly, a few risk-based confirmation items utilizing personal conduct standards are as of now being marketed [3, 4] and utilized chiefly in segments like the monetary administration's industry. A main Risk-based validation instrument as of now accessible is the RSA Adaptive Authentication System (RAAS) that ensures straightforwardly the online client action both at the login and transactional dimensions by examining misrepresentation pointers, client profiles, transaction behavioural pattern, etc [5]. In any case, a key distinction between these instruments is that notwithstanding customary information sources, for example, IP address and gadget attributes, ASS gathers and procedures keystroke elements biometrics (in spite of the fact that utilizing fixed content discovery). The absence of distributed trial results makes it be that as it may, difficult to evaluate the adequacy of the above business frameworks.

As far as anyone is concerned, a large portion of the distributed recommendations in the exploration writing has been at the transactional level [6].

In this specific situation, Dimmock et al. [6] presented a computational risk evaluation method for dynamic and adaptable access decision-making. The proposed methodology permits putting together access control choices with respect to risk and trust instead of on accreditations as it were. Be that as it may, the methodology expects earlier information of results of every conceivable blend of states and activities during the basic leadership process, which isn't sensible. Moreover, the subjects in their model are self-governing specialists, not people; and we realize that human personality and conduct are fundamental parts of risk-based verification.

A versatile verification instrument was proposed by Abu Bakar and Haron [7], which used a Unified Authentication Platform that incorporated multifaceted confirmation and Single sign-on. The model comprised of a trusted motor to produce examples and assess trust score. Trust score assessment depended on the verification technique quality, client login parameters and application security necessity. Shi et al. [1] connected a learning calculation to give the machine a chance to get familiar with the past behaviour of the client and create a client model. The probability of the client being certifiable is determined dependent on the client model and the as of late observed behaviour.

III. METHODOLOGY

Our proposed model has three blocks. User Metadata retrieval block, Risk calculation block and an adaptive authentication system block as depicted in figure 1.

A. User Metadata Block

At the point when the framework gets a solicitation for validation, the client enters his secret key and the verification server recovers the client parameters required to demonstrate client conduct. The client variables incorporate the accompanying. These parameters structure the contributions to the risk calculator input as appeared in figure2.

1. IP address
2. Location of the client
3. Time zone
4. Login Time
5. Operating System Version
6. Browser Version
7. Device Type
8. Number of Authentication Failure

Each client's past login records that fill in as the client's pattern profile as per [8] are put away at the confirmation server. When the client enters the right secret phrase, Risk Calculation Block is enacted. The as of late recovered relevant data of the client and client's past login records are made accessible to the Risk Calculation Block.

B. Risk Calculation Block

The Risk Calculation Block is the centre part of the Risk-based verification plot which can anticipate the risk dimension of a client. In this paper, the Risk Calculation Block is planned to utilize Autoencoder for AI.

The fundamental goal of the risk motor is to recognize inconsistencies in the information. For such purposes, labelled preparing set isn't fundamental. For unlabelled preparing dataset, a supervised learning algorithm can't be used. This requires the utilization of the unsupervised learning algorithm. Auto Encoder is one such unsupervised learning calculation. Subsequently, Auto Encoder is trained to utilize the information with just real client designs. It tends to be utilized when there is a lack of peculiarity information.

The output of Auto Encoder is set True or false, if it is able to regenerate the data- then it gives True or false otherwise. Hence, 'true' yield shows that the client is certifiable, while 'false' yield demonstrates that the client might be deceitful. In the event that the yield results in a bogus value, table 1 ought to be utilized to gauge the Risk score of the client. A risk score is determined to utilize Equation 1. A higher Risk score relates to a higher risk level.

Equation 1:

Risk score = Σuser_metadata_value* user_metadata_weight□□□

Where,

User_metadata_value=0

If behaviour exists in the past login records = 1

Otherwise.

User_metadata_weight is allotted as per table 1. These weights have been doled out in the wake of considering the effect every parameter would have in deciding potential risk.

TABLE 1. USER PARAMETER WEIGHTS

User parameters	Weight
Browser Version	1
Operation System Version	2
Login time	3
IP address	4
Device Type	5
No. of Authentication failure	6
Location of the client	7
Time zone	8

C. Adaptive Authentication System Block

The risk score/likelihood determined by risk calculation block is bolstered to the Risk Manager, which orders the risk level depends on the risk score/likelihood. The client is tested with a verification strategy like security questions, or OTP, or graphical password [9,10] related to each Risk level as referenced in table 2. Lower Risk levels have been related to shorted likelihood and Risk score runs as against the bigger limit ranges for higher risk levels. An effective login is stored as a certified record at the server, while a fruitless login endeavour is stored as a fake example.

TABLE 2 Authentication methods for each risk level

S No.	One-class AutoEncoder Risk Score (S)	Risk Level	Authentication Method
1	1≤S≤6	1	OTP token
2	7≤S≤18	2	Security Questions
3	19≤S≤29	3	Graphical Password
4	30≤S≤36	4	Digital Signature

IV. RESULTS AND DISCUSSION

The proposed method has been executed utilizing Python. For login, the client gives the username and password to his/her first login endeavour after enrolment. The machine at that point learns the client behaviour by dissecting the client parameters recovered at each login endeavour and raises the risk level at whatever point it experiences new client conduct.

The paper expects the accompanying limitations:
a) A change in login time is considered as a risk just if it surpasses an edge of two hours from the standard login time of the client.
Assume that the client's ongoing 10 exchanges are as specified in table3. The table indicates that the user has logged in utilizing similar parameters at various time interims and thus has been named a real client. During the beginning phases of client login, the risk calculation block stays inert. Once the chronicled dataset is populated with at least 10 records for each client, the risk calculation block is enacted to predict the risk dimension of the client. Table 3 demonstrates that the client has, for the most part, logged in from Ambala utilizing a windows 10.0 PC and chrome program at various occasions of the day.

Assume that the client attempts to log in with the user parameters as appeared in table 5 in the wake of having built up a behaviour profile.

Table 5 considers four unique situations for all the risk calculator system sorts. The main situation clarifies a circumstance where the client gets to the administration from an alternate area. The second situation compares to a state wherein the client signs in utilizing an alternate OS and a browser. The third arrangement of client parameters shows that the client has gotten to the administration from an alternate area utilizing distinctive OS and browser after 3 fizzled login endeavours. In the last circumstance, the client signs in from an alternate time zone with the various OS, browser after three ineffective login endeavours. The comparing AutoEncoder Risk Calculation system block yields show the likelihood of the client being a deceitful high-risk score of the client. The risk score is straightforwardly corresponding to the risk dimension of the client. Contingent upon the acquired values, the client is mentioned for further accreditations as referred to in table 2.
It is clear from table 5 that the risk score determined by AutoEncoder is more pertinent to the given situations, which seem extraordinary for all the model cases, for example, an adjustment in OS alone has brought about the least risk level, while an adjustment in time zone has come about in the highest risk level. The likelihood esteems for the first two scenarios demonstrate that the client is certifiable regardless of the change in area or change in OS/browser. The qualities created by AutoEncoder model are impressive despite the fact that a higher risk level would have been favoured for change in the area. Subsequently, AutoEncoder risk motor is progressively favoured for our test information. In any case, it can't be summed up that AutoEncoder calculation would dependably perform better in identifying irregularities as the presentation of any AI put together risk depends with respect to the training data.

To abridge the working of the Model, when the client's past login records, as appeared table 3, are encouraged to the Risk Calculation system block, it learns the standard conduct of the client and manufactures a model to mirror the client profile. In this way, when the client solicitations access with various parameters, the model predicts a higher risk level to abstain from disguising assaults. On the off chance that the client can log in in the wake of giving the additional

certifications, the machine records the new example as protected/safe. Else, the client's conduct is set apart as false.

A. HIGHLIGHTS AND ADVANTAGES OF THE PROPOSED METHOD

Regardless of the presence of a few ways to deal with multifaceted verification, relatively few examinations have been effectuated to decide choice technique for various authentication elements utilizing AI calculations. AI calculations offer the best way to show client conduct and finding the risk level related to every client. This paper utilizes calculations based on AutoEncoders to investigate client conduct.

Although Bayesian calculation has been utilized in the past to display client conduct [11], the model uses managed to learn calculation and along these lines can't give productive outcomes except if it is prepared to utilize certified and deceitful examples. This paper subsequently proposes the utilization of AutoEncoders calculation in circumstances where the two kinds of records may not be accessible. The risk score decided to utilize equation 1 relies upon the techniques for confirmation common in most multifaceted Authentication frameworks.

The proposed strategy likewise elevates ease of use notwithstanding security. A certified client isn't required to pass various elements of confirmations to demonstrate his authenticity, while just a suspicious client needs to pass all the verification strategies he is tested with. Therefore, clients are less vexed with the validation procedure.

TABLE 3 User Login Records

UID	IP address	Location	Time Zone	Login Time	OS	Browser	Mobile	Failed Attempts	Class
DBBC21A2	192.168.116.113	Pune	IST	9:11:44	Android Oreo	UC Browser Browser	Redmi Y2	0	Genuine
DBBC21A2	192.168.116.113	Pune	IST	11:24:31	Android Oreo	UC Browser Browser	Redmi Y2	0	Genuine
…..									
DBBC21A2	192.168.116.113	Pune	IST	22:53:13	Android Oreo	UC Browser Browser	Redmi Y2	0	Genuine

TABLE 4 Example for fraudulent patterns

UID	IP address	Location	Time Zone	Login Time	OS	Browser	Mobile	Failed Attempts	Class
DBBC21A2	103.5.19.11	Pune	IST	10:17:46	Android Marshmellow	UC Browser Browser	Redmi Y2	4	Fraudulent
DBBC21A2	103.5.19.11	Pune	IST	10:43:23	Android Marshmellow	Firefox	Redmi Y2	0	Fraudulent
DBBC21A2	103.5.19.2	Pune	IST	13:5:5	Android Marshmellow	Firefox	Redmi Y2	0	Fraudulent

TABLE 5 Four User login scenarios with Probability/Risk score

Scenario	I	II	III	IV
UID	DBBC21A2	DBBC21A2	DBBC21A2	DBBC21A2
IP address	1.22.247.55	192.168.116.113	1.22.247.55	192.154.1.11
Location	Ambala	Ambala	New Delhi	NewYork
Time Zone	IST	IST	IST	PST

Login Time	16:09:57	16:41:33	16:55:03	03:15:19
OS	Android Oreo	Android Marshmellow	Android Marshmellow	IOS
Browser	UC Browser Browser	Opera Mini	Opera Mini	Safari
Mobile	Redmi Y2	Redmi Y2	Redmi Y2	Iphone 6S
Failed Attempts	0	0	3	3
AutoEncoder	11	10	32	48

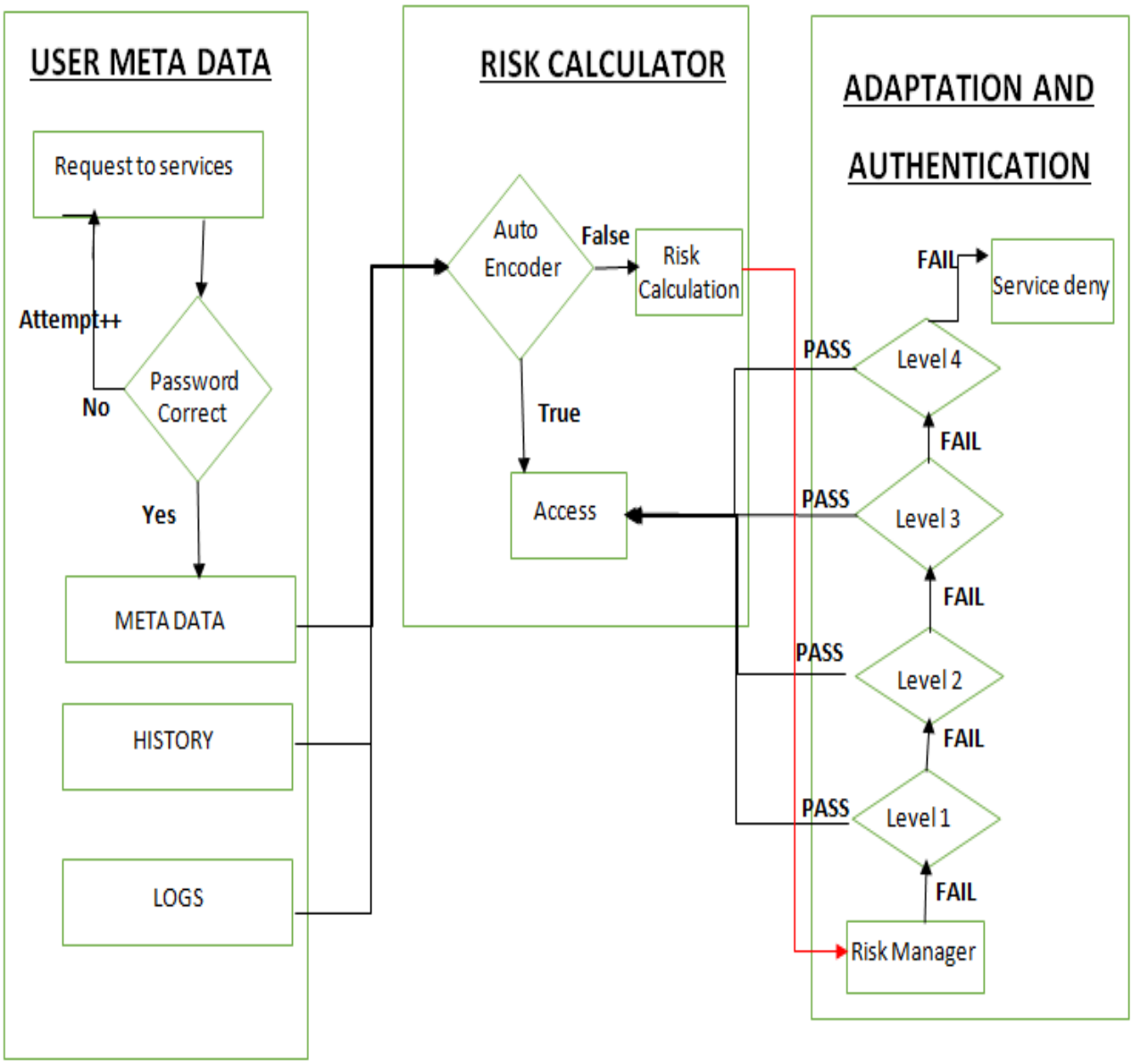

Figure 1: Architecture of Autoencoder

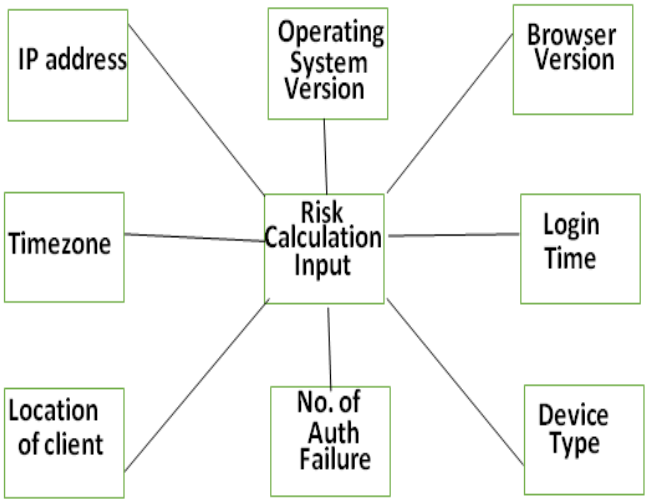

Figure 2: Risk in Authentication

V. CONCLUSION

Risk-Based Authentication offers further security to an online application. It thinks about different parameters before choosing whether or not to concede access to the user. Besides, the client is tested with additional variables of confirmation if his conduct seems suspicious of the risk calculation system block. Any variation in the login parameters is seen as a potential risk by the risk calculation system block and the framework may request an extra check from the client. Past login records of the client contain noteworthy ramifications in the structure of the risk calculation system block. Plus, the proposed strategy offers decisions for risk calculation system block to permit activity during circumstances where there is an absence of preparing records for false examples, requires the determination of risk using Auto Encoder. Also, as an extra proportion of security, the verification methodology is intended to be completed on the client's cell phone. Subsequently, a client's record can't be undermined except if the faker uses the genuine client's cell phone.

Four situations have been talked about in the paper to draw out the four plausible risk levels and the result of each risk has been looked at when an adjustment in the client conduct is experienced. The paper in this manner proposes a very secure technique to shield online records from digital dangers.

REFERENCES

[1] Kumar Abhishek, SahanaRoshan, Prabhat Kumar and Rajeev Ranjan. "A comprehensive study on Multifactor Authentication Schemes". Advances in Computing and Information Technology, 177, pp. 561-568,2013.

[2] Tubin G (2005) Emergence of risk-based authentication in online financial services: You Can't Hide Your Lyin' IP. Whitepaper #V43:15N, Tower Group, May

[3] Lian S, Chen X, Wang J (2012) Content distribution and copyright authentication based on combined indexing and watermarking. Multimedia Tools Appl 57(1):49–66

[4] Orozco M, Graydon M, Shirmohammadi S, El Saddik A (2012) Experiments in haptic-based authentication of humans, International Journal of Multimedia Tools and Applications - Springer Science + Business Media B.V.

[5] Issa Traore & Isaac Woungang & Mohammad S. Obaidat & Youssef Nakkabi & Iris La in Springer Science+Business Media New York 2013 DOI 10.1007/s11042-013-1518-5

[6] Dimmock N, Bacon J, Ingram D, Moody K (2005) Risk models for trust-based access control. In Proc. of the 3rd Annual Conference on Trust Management (iTrust'05),

[7] Mohan V. Pawar, Anuradha J. " Network Security and Types of Attacks in Network". Procedia Computer Science 48, pp. 503 – 506,2015.

[8] Oded Peer, Yedidya Dotan, Yael Villa and Marcelo Blatt. "USING BASELINE PROFILES IN ADAPTIVE AUTHENTICATION". US Patent 8,621,586 B1, December 31,2013.

[9] Misbahuddin, Dr Mohammed & Premchand, P & Govardhan, Dr. (2008). A user-friendly password authenticated key agreement For web based services. 10.1109/INNOVATIONS.2008.4781766.

[10] Environment", in International Conference on Advances in Computing, Communication and Control (ICAC3'09)

[11] DipankarDasgupta, Arunava Roy and Abhijit Nag. "Toward the design of adaptive selection strategies for multi-factor authentication". computers& security, pp. 85–116,2016.

Authors Profile

Mr Bharat Swaroop Sharma is pursuing Bachelor of Engineering in Information and Technology from Army Institute of Technology, Savitribai Phule Pune University, India. He is a CSI member since 2017. And Is an ML intern at BlueBricks Technologies India since June 2019.

Mr. Sidharth Singh is pursuing Bachelor of Technology in Computer Science from Maharishi Markandeshwar University, Ambala, India. He has done his training from IIT,Kanpur in Blockchain and is currently leader of Developer Student Club, MMDU (Sponsored by Google).

