ISSN: 2229-6948(ONLINE)
DOI: 10.21917/ijct.2018.0267

ICTACT JOURNAL ON COMMUNICATION TECHNOLOGY, SEPTEMBER 2018, VOLUME: 09, ISSUE: 03

DIGITAL IDENTITY MODELLING FOR DIGITAL FINANCIAL SERVICES IN ZAMBIA

Wakwinji Inambao[1], Jackson Phiri[2] and Douglas Kunda[3]
[1,2]Department of Computer Science, Mulungushi University, Zambia
[3]Department of Computer science, University of Zambia, Zambia

Abstract

Identification and verification have always been at the heart of financial services and payments, which is even more the case in the digital age. So, while banks have long been trusted to keep money safe, is there a new role for them as stewards of digital identity? Governments should, in consultation with the private sector, develop a national identity strategy based on a federated-style model in which public and private sector identity providers would compete to supply trusted digital identities to individuals and businesses. Back then, when the world seemed smaller, slower and more local, physical identity documents were adequate for face-to-face transactions. However, the Internet changed everything. It shrank distances, created new business models and generally sped everything up. From the innovation lifecycle to access to information, processes and the clock-speed on risk, the Internet has accelerated everything. The use of Internet in doing business has grown over the years in Africa and Zambia in particular. As such, the incidences of online identity theft have grown too. Identity theft is becoming a prevalent and increasing problem in Zambia. An identity thief only requires certain identity information to decimate a victim's life and credit. This research proposes to identify and extract various forms of identity attributes from various sources used in the physical and cyberspace to identity users accessing the financial services through extracting identity attributes from the various forms of identity credentials and application forms. Finally, design a digital identity model based on Shannon's Information theory and Euclidean metric based Euclidean Distance Geometry (EDG) to be used for quantifying, implementation and validating of extracted identity attributes from various forms of identity credentials and application forms, in an effective way.

Keywords:

Identity Theft, Identity Credential Attributes, Security, Digital Identity Model, Business

1. INTRODUCTION

Identity theft involves acquiring key pieces of someone's identifying information, such as name, date of birth, social security number and mother's maiden name, in order to impersonate them [1]. This information enables the identity thief to commit numerous forms of fraud which include, but are not limited to, taking over the victim's financial accounts, opening new bank accounts, purchasing automobiles, applying for loans, credit cards, and social security benefits, renting apartments, and establishing services with utility and phone companies [4]. The authors in [12] defined digital identity as "a virtual representation of a real identity that can be used in electronic interactions with …". The authors [3] argued that each year, millions of Americans discover that a criminal has fraudulently used their personal information to obtain goods and services and that they have become victims of identity theft. According to [4], under federal law, identity theft occurs when someone uses or attempts to use the sensitive personal information of another person to commit fraud. A wide range of sensitive personal information can be used to commit identity theft, including a person's name, address, date of birth, Social Security number (SSN), driver's license number, credit card and bank account numbers, phone numbers, and even biometric data like fingerprints and iris scans [1].

The authors in [37] reported that, "*The threat from computer crime and other information security breaches continues unabated and the financial toll is mounting……*"

Information technology security (Cybersecurity), focuses on protecting computers, networks, programs, and data from unintended or unauthorized access, change, or destruction [25].

According to [14], government agencies, the military, corporations, financial institutions, hospitals, and other groups collect, process, and store a great deal of confidential information on computers and transmit that data across networks to other computers. With the growing volume and sophistication of cyber-attacks, ongoing attention is required to protect sensitive business and personal information, as well as safeguard national security. African countries are more worried due to undeveloped technology facilities.

Cyber security involves protecting information and systems from major cyber threats, such as cyber terrorism, cyber warfare, and cyber espionage [26]. In their most disruptive form, cyber threats take aim at secret, political, military, or infrastructural assets of a nation, or its people [25]. Cyber security is therefore a critical part of any governments' security strategy [4].

According to [4], cyber terrorism is the disruptive use of information technology by terrorist groups to further their ideological or political agenda. This takes the form of attacks on networks, computer systems, and telecommunication infrastructures [1].

Cyber warfare involves nation-states using information technology to penetrate another nation's networks to cause damage or disruption [55]. The authors in [55] argued that cyber warfare attacks are primarily executed by hackers who are well trained in exploiting the intricacies of computer networks and operate under the auspices and support of the nation-states. Rather than "shutting down" a target's key networks, a cyber-warfare attack may intrude networks for the purpose of compromising valuable data, degrading communications, impairing infrastructural services such as transportation and medical services, or interrupting commerce [6].

If Identity theft involves acquiring key pieces of someone's identifying information as pointed out by [1], cyber espionage then entails, the use of information technology to obtain secret information without permission from its owners or holders. According to [20], cyber espionage is most often used to gain

strategic, economic, political, or military advantage. It is conducted through the use of cracking techniques and malware. In the US, according to [9], the Office of the National Counter Intelligence Executive released a report in 2011 officially acknowledging the legitimate threat of cyber espionage and its potential to damage the United States' strategic economic advantage. The purpose of this study is to improve the authentication in the digital identity management system and particularly improve the security in digital finance services. The use of mathematical models in the quantification of identity attributes in this paper is hoped to enhance security values to different identity attributes used in financial services and later on to other services.

2. DIGITAL FINANCIAL SERVICES

The authors in [2] argued that Digital Financial Services have significant potential to provide a range of affordable, convenient and secure banking services to the poor people in the developing nations like Zambia. According to estimates [2], more than 400 million people are linked globally through basic mobile payments services, allowing them to send money, pay bills, or purchase prepaid electricity with greater ease, affordability and access. Increasingly governments are adopting digital finance to deliver social safety net cash payments and trying to make collection of fees/tariffs more efficient [2]. Yet there remains a long way to go in digital finance. The authors in [2] pointed out that ecosystems take time to develop before one-quarter to a half of adults begin to use basic payment services. However, there are also barriers that hinder the progression from payments to solutions "beyond payments". For example, many payments services remain relatively closed, making the integration of a broader range of solutions into existing payment platforms costly and cumbersome. In addition, many financial services such as savings and loans require significant physical touch points between customers and providers, making them difficult to scale [12].

Digital financial services (DFS) can expand the delivery of basic financial services to the poor through innovative technologies like mobile-phone-enabled solutions, electronic money models and digital payment platforms [2]. The authors in [1] argued that digital channels can drastically drive down costs for customers and service providers, opening the door to remote and underserved populations [2]. Financial regulators around the world have realized the tremendous role DFS can play for financial inclusion and seek to unlock this potential by creating enabling environments for digital financial services [2] [1].

2.1 THE STATE OF DIGITAL FINANCIAL SERVICES (DFS) IN ZAMBIA

According to [1], Zambian DFS agents have the same volume of transactions as other countries but making the least revenue. In their report, [1] argued that albeit there is high volume of transactions, the low revenue is due to the low value transactions which bring about very small commissions for agents.

The authors in [3] reported that Bank of Zambia noted high transaction fees levied by various financial service providers continue to affect the growth of the services. In the report, the bank was quoted saying, "high levels of agent exclusivity put a lot of pressure on the agents to achieve a critical mass of customers and transactions". The bank argued that financial service providers should adopt a shared agent model, a move that would increase the revenue and most likely the motivation for agents. This, the reporter viewed as important for the agents in rural areas. However, Lack of awareness of DFS products that are available on the market also poses a challenge [2]. There is need to increase awareness of DFS and service providers should use several channels to market their products.

According to [1], Zambia has high levels of agent exclusivity (91 percent) compared to other countries in Africa. There is a need to watch out for cybercrime while ensuring that DFSs play a big role towards the attainment of the targets set out in the National Financial Inclusion Strategy. New risks such as cybercrime that come with the digital era, will have to be mitigated and monitored [2]. However, there is need to ensure that regulation to mitigate the risk is proportionate to the risk as over regulation may stifle innovation. Thus, there is need to strike an appropriate balance.

3. RELATED WORKS

The authors in [29] built up a mathematical model of different core topics covering a wide range of vocabularies related to Identity Management. This was done by building up a mathematical model of Digital Identity that was used to analyze different aspects of Identity Management. It suffices to note that the mathematical model was also used to develop a solid understanding on different topics of Identity Management. However, an effort was undertaken to develop an extensive list of vocabularies [29]. The aim of the model was to build up a common language to remove inconsistencies. According to [29], the number of online services and users that access service has led to an increase in digital identifiers. It is argued that, [29] the issues to manage the identifiers and their corresponding credentials has become a difficult responsibility. The authors in [30] proposed an identity model to facilitate management of online identities [29] [14].

In [29]'s research, a mathematical framework was developed to model the main aspects of identity management and used the framework to model many aspects of identity management systems.

On the other hand, [4] proposed an algorithm in which a pipe-lined system was designed to take identity theft news stories from the Internet as input and generates the analytics that help [4] better understand the identity theft process as a result. The authors in [4] developed the model in such a way that data collected through the process of getting the news articles was linked from the Internet. It is argued by [4] that the Identity Theft Resource Centre Additionally, links are hauled out from openly available annual identity theft reports on the Identity Theft Resource Centre. The breach report given by [4] consists of data breaches collected from an assortment of media sources.

To manage digital identity information [12] proposed Digital Identity Management System model as a solution. In-line with the solution of managing digital identity information, [12] outlined artificial intelligence and biometrics on the current unsecured networks to maintain the security and privacy of users as well as the service providers. The authors in [12] also proposed the use of the multimode neural network which was outlined as having ten neurons and three layers. The authors displayed the Neuron 1 to 5

which was used as input neurons while neuron 6, 7, 8 and 9 were depicted to represent the groupings of; physical metrics, biometrics, pseudo metrics and device metrics in that order. On the other hand, [12] let Neuron 10 to be the output neuron with y_{10} used for multimode authentication. However, the attributes which drew their strengths from the metrics were used as input neurons for the network. The authors in [12] allowed the input variables x_1 to x_2 to form the vector for the physical metrics whereas the variable x_3 formed the vector for biometrics. The authors in [12] designed Neurons x_4 to form the pseudo metrics vector with neurons x_5 to form the device metrics vector.

Wherefore, [12] used propagation algorithm for dealing with initial weights w assigned at random between the values of 0 and 1 the authors represented the "activation function as;

$$X = \sum_{i=1}^{n} x_i w_i - \theta$$

With X being the net weighted input to the neuron, n is the number of neuron inputs, and θ is the threshold to the neuron. x_i is the value or strength of the input variable and w_i the its respective weight" [12].

The authors in [26] developed an identity management model for automated processing of identity information between distributed ecosystem partners. The model was based on the new OASIS SAML standard to provide interoperability and convergence between existing identity technologies [26]. The author presented a basic and extended identity models for single services and service compositions. [26] wanted to allow SMEs to use and enhance their current identity technology with a practical and easy to implement identity management solution that scales up to the dynamic and distributed nature of digital ecosystems. In addition, [26] aimed to automate the process of identification between ecosystem partners. The emphasis of the model was to find practical solutions which are clear and easy to implement.

However, [26] proposed the identity management model for decentralized peer-to-peer ecosystem domains. In this model, all users were considered equal and there was no hierarchy of ecosystems. Nevertheless, any peer could be considered a Credential Provider or a Service Provider, or both. The authors in [20] outlined disadvantages concerning privacy or identity theft. As such [20] developed hybrid models that linked digital identities generated by service providers with the identities that users provide. The authors in [20] argued that the emergence of federated identity management gave rise to complex scenarios in which identity management is carried out in a fragmented and adaptable way. In agreeing with [10], [20] argued that fragmentation meant that whoever issues and validates the credentials of a digital identity does not necessarily have to be the owner of the resource. In other word, [20] deduced that the model was able to provide identity, as well as its management, as a service (IDaaS).

According to [10], cloud-based service method for user identity attribute service in cyberspace, and the method based on the service model of multilevel cyber identity management provides user identity attribute service for cyber applications [12]. The authors in [10] pointed out that, security of web applications along with user privacy is protected by providing fine-grained access control with user identity attribute and strong authentication services [26]. And so, [10] argued that it is necessary to explore the new cyber user identity attribute service method which include the multilevel privacy protection, providing fine-grained access control with user identity attribute as well as protecting user privacy under assuring cyber application security [12] [14].

On the other hand, [8] proposed a novel framework for formalizing and comparing identity-related properties. The framework that [8] designed employed the notions of detectability, associability and provability to assess the knowledge of an adversary. The author [8] indicated how the established notions used to specify known identity-related properties and classify them with respect to their logical relations and privacy strength. The authors in [8] further demonstrate that the proposed framework is able to capture and compare several existing definitions of identity-related properties.

The personal information (PI) model, was designed [8] for use to represent the context of a computer system containing personally identifiable information.

In [40]'s study, key requirements for online transactions, and explores the legal challenges that most businesses must address to reap the benefits of federal identity were examined. In doing this, [40] focused on the emerging federated approach to identity, where an enterprise engages in online transactions in reliance on identity credentials issued by any one of several third parties, and individual can use the same identity credential to engage in transaction with multiple organizations. Focused on the federal identity approach which examined the need for operating rules to ensure that such an identity system functions properly in a way reckoned trustworthy by the businesses [40].

The authors in [40] argued that a federal identity system combines business, organizational, and technical functions and capabilities with legal infrastructure to expedite the collection, verification, communication, and authentication of identity information by and/or various parties [56].

4. METHODS OF IDENTIFICATION, MINING/ EXTRACTION, AND MODELLING OF CREDENTIAL ATTRIBUTES

4.1 IDENTIFICATION OF FINANCIAL CREDENTIAL TOKENS

In order to open an account at the financial institution, the applicant is required to fill in the application form which contains fields that are necessary for the identification of the applicant. Different types of forms of identity credentials used in identification of individuals around the world and Zambia in particular are identified [6] [12]. To be able to maximize on the collection of identity credentials, the use of questionnaires was employed. In the questionnaire, the types of credentials tokens known or used in financial institutions were solicited [55] as shown in Table.1.

For identification proof of the applicant, the following albeit not limited to, Birth certificates, Driver's license, Marriage certificates, National registration Card, Credit Card, Passport, voter's card, any photo ID card issued by the central government accompany the application forms. Unlike [6] who collected data from the websites, the researcher collected the credential tokens

or application forms of many kinds and used them to extract data [12] [14] [50].

4.2 IDENTIFICATION OF GOVERNMENT CREDENTIAL TOKENS

Being a citizen of a country, one needs to satisfy certain obligations which go with the acquisition of say, the birth certificates, national Registration cards, passports, voter's cards, to mention but a few. It is the obligation of every government to provide its citizens with national credentials which in a way go with rights and benefits of belonging. To acquire these documents, application forms have to be filled in. The application forms help government agencies to capture vital identity credentials of its applicants [13].

On the other hand, every nation has an obligation to provide documentations to other countries' citizens that want to get access to its land and services such as schools, employment. In most countries, a foreign national has to apply for a visa to be allowed in the country [6]. There are various application forms as well which need to be filled in. The application forms contain vital identification attributes that a country may require the non-citizens to fill in. The identity attributes in the application forms are for use in the identification of the foreign national, with periods of stay and other issues or vital identity issues [14].

4.3 MINING AND EXTRACTION OF CREDENTIAL ATTRIBUTES

Credential tokens used for application of services discussed earlier contain credential attributes. Having identified all the credential tokens of financial and governments, the vital and compulsory attributes from the credential tokens are extracted from the forms. The data mining tools used in the process of extracting the credential attributes from the credential tokens or application forms are antconc 3.2.1w, Concapp and Texttat-2.

4.4 MODELLING OF CREDENTIAL ATTRIBUTES

After credential attributes were extracted from the credential tokens, the attributes were then quantified. Shannon's information theory was used to quantify the attributes. In this process, the attributes are weighted according to how prominent the attributes are in the credential tokens [14] [15]. In making sure that the digital identities are varied according to how sensitive or prominence the attributes are in the credential token, Shannon's entropy helps to weigh the attributes.

Euclidean Geometric Distance was used to assign ranges to credential attributes that have been weighed by Shannon's information theory by comparison on the basis of the number of attributes in each credential token.

4.5 IDENTIFICATION OF THE IDENTITY CREDENTIAL TOKENS

The use of questionnaire was employed to respondents in two towns of Lusaka and Kabwe which were selected randomly from the ten provincial centres to solicit for a list of identity documents which are used in financial institutions when applying for banking services. More also Kabwe and Lusaka were selected for this activity for the following reasons: First, Lusaka is the capital city of Zambia, where the passport and national registration offices are well equipped with information and facilities, it was appropriate to get information from the officers there. Secondly, the researcher is based in Kabwe and near Lusaka, making it easy for the study. In short, the selected towns are convenient for the study and the researcher thereof. In addition, the respondents were required to list the accompanying identity credentials used as identity ID (government identity credentials) [49]. The Table.1 shows common identity credentials identified by respondents used for various services. However, each service has specific identity credentials used.

Table.1. Identity Credentials used for various services

Identity credentials	Required
Passport	Need to identify Identity credentials required strictly for accessing financial services out of the listed and many more to be identified.
Clinic card	
National Identity card	
Driving licence	
Bank card	
Student card	
Voter's card	
Baptism card	
Education card	
Social security card	
Insurance card	
Credit card	
Hospital card	
Military card	
Marriage certificate	
Fire aim certificate	
Pet corticated	
Adoption certificate	
Payslip	
TPIN certificate	

5. PROPOSED MODEL

This section presents the model used in the identification, extraction and quantification of identity credential attributes indicated in Fig.1. The identity credential tokens were subjected to data mining tools to extract identity credential attributes. The extracted attributes were then weighted by Shannon's Information theory after being graded by the respondents on uniqueness, consistency, verifiability and persistency captured through the questionnaire.

The Shannon information content of an outcome $x = a_i$ is defined as,

$$h(x = a_i) = \log_2 \frac{1}{P(x = a_i)} \quad (1)$$

With the event $x = a_i$ depicting uniqueness, consistency, verifiability and persistence and $P(x=a_i)$ being the probability of each of the listed Internet Identity properties above [14].

However, in our results, the entropy of the ensemble of $x = a_i$ is defined as the Shannon's Information content of the outcome:

$$H(X = a_i) = \sum P(x = a_i) \log_2 \frac{1}{P(x = a_i)} \quad (2)$$

With the event $x = a_i$ depicting uniqueness, consistency, verifiability and persistence and $P(x=a_i)$ being the probability as above [12].

The standardized Euclidean distance between Internet Identity properties of the Identity attributes x_j and y_j indicated in Table.1 is defined as:

$$d_{xy} = \sqrt{\sum_{j=1}^{J} \left(\frac{x_j}{s_j} - \frac{y_j}{s_j} \right)^2} \quad (3)$$

where s_j is the standard deviation of the Internet identity properties of the identity attributes. Now Eq.(3) can be written as:

$$d_{xy} = \sqrt{\sum_{j=1}^{J} w_j \left(x_j - y_j \right)^2} \quad (4)$$

where $w_j = \frac{1}{s_j^2}$.

In this paper, after standardizing the scores, Eq.(4) became

$$d_{xy} = \sqrt{(x_1 - x_2)^2 + (x_3 - x_4)^2 + \ldots + (x_n - x_{n+1})^2} \quad (5)$$

After Shannon's Information theory was used in the weighting, Euclidean Geometric Distance was then applied to the weights to determine the ranges and prominence in the values.

These values should determine the digital identity key to identification credentials. The Fig.1 shows the proposed model for identification of identity credentials, extraction of the identity credential attributes and quantification of the attributes.

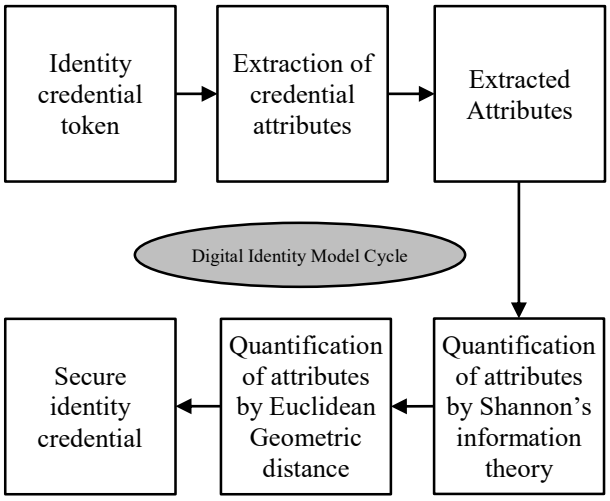

Fig.1. Proposed model for identity credential attributes extraction and quantification

In this study, financial services credential tokens are used as the sources of the identity attributes. All forms (documents) that were in the PDF format were converted into text format and then used AntConc 2.5.7 to extract the attributes. The conversion from PDF to text was done with the help of an online application [57].

A total of 45 documents (application forms) used in financial services were used in this study to create word, frequency, collected frequencies and frequency the document form. The Fig.2 shows a total of 45 documents (application forms) in the corpus with 4848 concordance hits. The Fig.3 shows the hits that some documents received, such as application form word.txt had 210 hits, bank accounts.txt had 100 hits to mention but two documents.

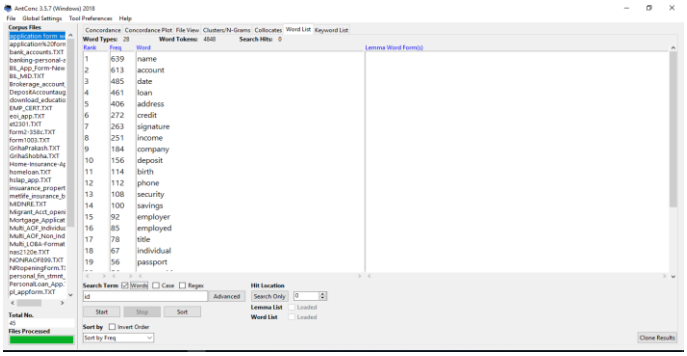

Fig.2. AntConc 2.5.7 results after extraction from the application forms

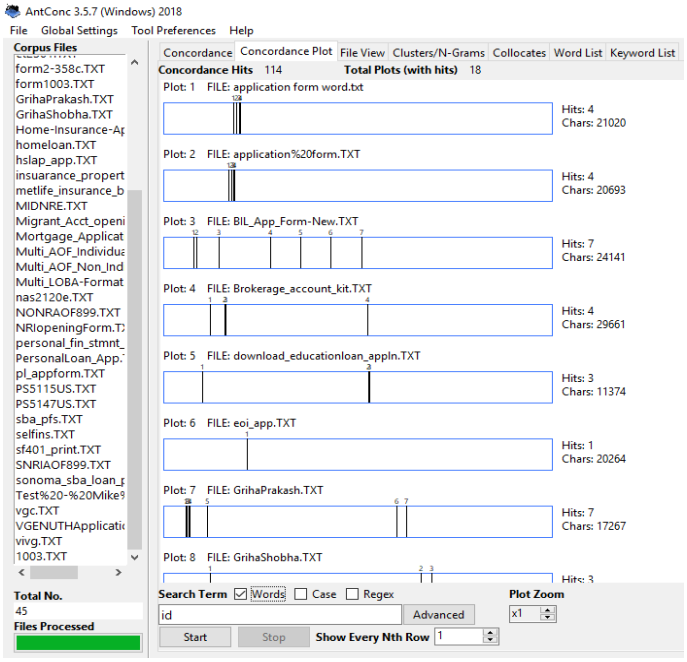

Fig.3. AntConc corpus search results for names showing terms and document frequencies

As in [12], questionnaires were administered to 36 respondents and asked to rank attributes out of five using the properties used by [13]. In the ranking, for instance, for consistent, a_1 depicted not consistent and a_5 depicted very consistent. The average scores obtained from the questionnaires are shown in Table.2.

The average scores were then converted into probabilities before Shannon's information theory was applied as shown in Table.2. To calculate the weighted values of the identity attributes shown in Table.3, the Shannon information theory defined in Eq.(1) and Eq.(2) was applied to the probability average scores as indicated in Table.3.

With results obtained from the questionnaires in Table.2, the scores were standardized to give the results in Table.4. Using standardized scores, the Euclidean distance was used to calculate the aggregated squared differences in standard deviation units of each variable. In Table.5 only ten components are show out of the thirty-six components obtained. It must be noted that only ten out of thirty-six components are shown in this paper.

$$\text{Using } d = \sqrt{(x_1 - x_2)^2 + (x_3 - x_4)^2 + \ldots + (x_n - x_{n+1})^2} \quad (6)$$

Finally, the study obtained a Standardized Euclidean distances between the 36 scores, based on the five aspects: Uniqueness, consistence, verification, persistency and trust obtained from the questionnaires, showing part of the triangular distance matrix in Table.5 [14] which emanate from the standardized scores of the ranked attributes shown in Table.4.

Table.2. Ranking attributes of Internet properties

Attributes	**Average Scores of properties of Internet Identities**				
	Uniqueness	**Verifiability**	**Consistency**	**Persistent**	**Trust**
National ID	3.86	3.53	3.75	3.44	4.47
Address	2.64	2.86	2.92	2.50	3.06
Date of birth	2.06	3.39	3.11	3.67	3.28
Gender	1.67	3.33	2.94	3.44	3.06
Security (pin)	4.39	4.39	4.64	4.58	5.00
Signature	5.00	4.00	3.75	3.39	3.94
Email	3.64	3.42	3.58	4.19	1.53
First name	2.92	3.42	3.19	2.47	2.56
Surname	4.25	3.58	4.53	3.31	3.64
Sex	1.64	2.92	2.72	2.19	2.17
Height	3.69	3.31	3.61	3.14	4.14
Phone ID	4.47	3.50	3.61	2.72	3.50
Race	1.53	3.28	3.25	4.25	4.08
Mother's name	3.78	2.50	3.61	2.33	4.36
Town of birth	2.42	3.47	3.36	2.53	2.67
Country of birth	3.33	2.81	3.83	3.33	3.03
Fingerprint	4.50	4.31	4.06	4.33	4.47
Weight	2.28	2.72	2.92	2.75	2.83
Eyescan	4.14	3.42	4.03	4.14	4.39

Table.3. Weighted values of the identity attributes

Attributes	$P(a_1)\log_2 \frac{1}{P(a_1)}$					$P(a_1)\log_2 \frac{1}{P(a_1)}$
	Uniqueness	**Verifiability**	**Consistency**	**Persistent**	**Trust**	
National ID	0.2489	0.2301	0.2318	0.2299	0.2627	1.204
Address	0.1934	0.2001	0.1960	0.1853	0.2049	0.980
Date of birth	0.1626	0.2242	0.2048	0.2395	0.2148	1.046
Gender	0.1399	0.2217	0.1973	0.2299	0.2049	0.994
Security (pin)	0.2699	0.2648	0.2657	0.2758	0.2816	1.358
Signature	0.2924	0.2497	0.2318	0.2275	0.2425	1.244
Email	0.2396	0.2254	0.2250	0.2610	0.1255	1.076
First name	0.2070	0.2254	0.2085	0.1839	0.1813	1.006
Surname	0.2645	0.2325	0.2617	0.2238	0.2301	1.213
Sex	0.1382	0.2028	0.1870	0.1692	0.1615	0.859

Height	0.2420	0.2205	0.2262	0.2162	0.2501	1.155
Phone ID	0.2731	0.2290	0.2262	0.1964	0.2243	1.149
Race	0.1314	0.2193	0.2109	0.2631	0.2480	1.073
Mother's name	0.2455	0.1825	0.2262	0.1766	0.2586	1.089
Town of birth	0.1821	0.2278	0.2157	0.1867	0.1867	0.999
Country of birth	0.2263	0.1975	0.2352	0.2250	0.2036	1.088
Fingerprint	0.2741	0.2616	0.2439	0.2664	0.2627	1.309
Weight	0.1747	0.1935	0.1960	0.1978	0.1946	0.957
Eyescan	0.2602	0.2254	0.2429	0.2588	0.2596	1.247

Table.4. Standardized scores of the ranked attributes using Internet properties given by respondents as shown in Table.3

Component	**Uniqueness**	**Consistence**	**Verification**	**Persistency**	**Trust**
s_1	0.2064	-0.8184	-1.3725	2.0421	-2.1442
s_2	-1.2795	-0.8184	-1.3725	-0.5835	-0.6877
s_3	1.6923	0.7322	0.4575	-0.5835	-0.6877
s_4	-1.2795	-0.8184	0.4575	0.7293	-0.6877
s_5	0.2064	-0.8184	0.4575	-0.5835	-0.6877
s_6	0.2064	0.7322	0.4575	-0.5835	-0.6877
s_7	0.2064	-0.8184	-1.3725	0.7293	0.7687
s_8	0.2064	2.2829	-1.3725	0.7293	0.7687
s_9	0.2064	0.7322	0.4575	0.7293	0.7687
s_{10}	-1.2795	-0.8184	0.4575	0.7293	0.7687

Table.5. Standardized Euclidean distances between the 10 samples, based on the five Internet properties given by respondents obtained in Table.3

	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}
s_2	3.35001								
s_3	3.818984	3.818984							
s_4	3.599889	3.599889	3.599889						
s_5	1.982736	1.982736	1.982736	1.982736					
s_6	1.55063	1.55063	1.55063	1.55063	1.55063				
s_7	3.098033	3.098033	3.098033	3.098033	3.098033	3.098033			
s_8	3.10126	3.10126	3.10126	3.10126	3.10126	3.10126	3.10126		
s_9	2.398602	2.398602	2.398602	2.398602	2.398602	2.398602	2.398602	2.398602	
s_{10}	2.147624	2.147624	2.147624	2.147624	2.147624	2.147624	1.55063	2.147624	2.147624

6. CONCLUSION AND FUTURE WORK

The study brought out a lot of vital points concerning identification and extraction of various forms of identity credentials, and quantification, implementation and validation of extracted identity attributes from application forms. There viewed literature also gave an understanding of the use of mathematical models in attaching values to identity attributes for the sake of protection and secrecy. The whole process of mathematical modelling was reviewed and in the process helped to come up with an identity model based on Shannon's Information theory and Euclidean metric based on Euclidean Distance Geometry used for quantifying, implementing and validating of extracted identity attributes from application forms.

It suffices to mention here that the baseline results on the ranking of identity attributes using the properties of Internet identities (uniqueness, verifiability, consistency, persistency and trust) were subjected to Shannon's Information theory and Euclidean Distance Geometry as discussed in details in the methods. The results of the Euclidean Distance Geometry validated those obtained by Shannon's Information theory. This study helps improve the authentication in digital identity

management system and particularly improve the security in digital financial services.

REFERENCES

[1] A. Soleimani, B.N. Araabi and K. Fouladi, "Deep Multitask Metric Learning for Offline Signature Verification", *Pattern Recognition Letter*, Vol. 80, pp. 84-90, 2016.

[2] Alliance for Financial Inclusion, "Mobile Financial Services: Mobile-Enabled Cross-Border Payments," Available at: https://www.afi-global.org/sites/default/files/publications/mfswg_guideline_note_no_14_en9-2.pdf.

[3] K. Nyati, "Mobile Money Accelerating Digital Financial Inclusion", Available at: http://www.daily-mail.co.zm/mobile-money-accelerating-digital-financial-inclusion/.

[4] Y. Yang, M. Manoharan and S.K. Barber, "Modelling and Analysis of Identity Threat Behaviors Through Text Mining of Identity Theft Stories", *Proceedings of IEEE Joint Intelligence and Security Informatics Conference*, pp. 24-26, 2015.

[5] E.B. Berroukech, E.B.Y. El Idrissi, R. Ajhoun and H. Lamrani, "Identity Management Systems: Laws of Identity for Models0 Evaluation", *Proceedings of 4th IEEE International Colloquium on Information Science and Technology*, Vol. 16, No. 6, pp. 746-740, 2016.

[6] D. Giles, "Constructing Identities in Cyberspace: The Case of Eating Disorders", *British Journal of Social Psychology*, Vol. 45, No. 21, pp. 463-477, 2007.

[7] D. Berbecaru and A. Lioy, "On the Design, Implementation and Integration of an Attribute Provider in the Pan-European eID Infrastructure", *Proceedings of IEEE Symposium on Computers and Communication*, pp. 27-31, 2016.

[8] M. Veeningen, B. De Weger and N. Zanno, "Modeling Identity-Related Properties and their Privacy Strength", Available at: https://security1.win.tue.nl/~zannone/publication/veen-dewe-zann-10-FAST.pdf.

[9] E.A. Fischer, "Cybersecurity Issues and Challenges: In Brief", Available at: https://fas.org/sgp/crs/misc/R43831.pdf.

[10] X. Zoua, B. Chen and B. Jina, "Cloud-based Identity Attribute Service with Privacy Protection in Cyberspace", *Proceedings of International Workshop on Information and Electronics Engineering*, 2012.

[11] W.T. Chee, C.L. Chen and M. Govindarasu, "Vulnerability Assessment of Cybersecurity for SCADA Systems", *IEEE Transactions on Power Systems*, Vol. 23, No. 4, pp. 1836-1846, 2008.

[12] Jackson Phiri, Tiejun Zhao and Jameson Mbale, "Identity Attributes Mining, Metrics Composition and Information Fusion Implementation using Fuzzy Inference System", *Journal of Software*, Vol. 6, No. 6, pp. 1025-1033, 2011.

[13] J.Phiri and Tiejun Zhao, "Identity Attributes Quantitative Analysis and the Development of a Metrics Model using Text Mining Techniques and Information Theory", *Proceedings of IEEE International Conference on Information Theory and Information Security*, pp. 390-393, 2010.

[14] Jackson Phiri, Tie-Jun Zhao and Johnson I. Agbinya, "Biometrics Device Metrics and Pseudo METRICS in a Multifactor Authentication with Artificial Intelligence", *Proceedings of 6th International Conference on Broadband and Biomedical Communications*, pp. 157-162, 2011.

[15] C. Kabuya, J. Phiri, T. Zhao and Y. Zhang, "Metric Based Technique in Multi-factor Authentication System with Artificial Intelligence Technologies", *Proceedings of International Conference on Future Wireless Networks and Information Systems*, pp. 89-97, 2012.

[16] L. Hansen and H. Nissenbaum, "Digital Disaster, Cyber Security, and Copenhagen School", *International Studies Quarterly*, Vol. 53, pp. 1155-1174, 2009.

[17] Z.D.O. Immigration, "E-Visa", Available at: http://www.zambiaimmigration.gov.zm.

[18] T.R. Society, "Progress and Research in Cybersecurity Supporting a Resilient and Trustworthy System for the UK, Available at: https://royalsociety.org/~/media/policy/projects/cybersecurity-research/cybersecurity-research-report.pdf.

[19] R.A. Kemmerer, "Cybersecurity", *Proceedings of 25th International Conference on Software Engineering*, pp. 1-11, 2003.

[20] L. Dandurand and S.O. Serrano, "Towards Improved Cyber Security Information Sharing", *Proceedings of 5th International Conference on Cyber Conflict*, pp. 1-16, 2013.

[21] N.J.D.O.C. Affairs, "The Cybersecurity Handbook", Available at:https://www.njconsumeraffairs.gov/News/Brochures/Cyber-Security-Handbook.pdf.\

[22] A.T.W. Paper, "New Paradigms of Digital Identity: Authentication and Authorization as a Service (AuthaaS)", Available at: https://www.elevenpaths.com/wp-content/uploads/2015/10/Telefonica_LVTI2N.pdf.

[23] R. Derakhshani and A. Ross, "Fast Automatic Retinal Vessel Segmentation and Vascular Landmarks Extraction Method for biometric Applications", *Proceedings of International Conference on Biometrics, Identity and Security*, pp. 1-7, 2010.

[24] W. T. Chee, C. L. Chen and M. Govindarasu, "Cybersecurity for Critical Infrastructures: Attack and Defense Modeling", *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, pp. 853-865, 2010.

[25] L. Huffman, "Exploring a National Cybersecurity Exercise for University", *IEEE Security and Privacy*, Vol. 3, No. 5, pp. 27-33, 2012.

[26] A.E. Joshi, "Student Centric Design for Cybersecurity Knowledge Empowerment", *Proceedings of IEEE International Conference on Technology Enhanced Education*, pp. 1-4, 2012.

[27] S. Cheang, "Conceptual Model for Cyber-security Readiness Assessment for Public Institutions in Developing Countries: Cabodia", *Proceedings of 4th International Conference on Computer Sciences and Convergence Information Technology*, pp. 1411-1418, 2009.

[28] H. Koshutanski, M. Ion and L. Telesca, "Distributed Identity Management Model for Digital Ecosystems", *Proceedings of International Conference on Emerging Security Information, Systems and Technologies*, pp 1223-1229, 2007.

[29] D.J. MacKay, "Information Theory, Inference, and Learning Algorithms", Available at: http://www.inference.org.uk/itprnn/book.pdf.

[30] N. Kshetri, "Indian's Cybersecurity Landscape: The role of Private Sector and Public-Private Partnership", *IEEE Computer and Reliability Society*, Vol. 13, No. 3, pp. 1-8, 2015.

[31] S.M. Ferdous, G. Norman and R. Poet, "Mathematical Modelling of Identity, Identity Management and Other Related Topics", *Proceedings of 7th International Conference on Security of Information and Networks*, pp. 1-5, 2011.

[32] S.M. Ferdous, G. Norman and R. Poet, "Mathematical Modelling of Identity, Identity Management and Other Related Topics", *Proceedings of 7th International Conference on Security of Information and Networks*, pp. 1-6, 2014.

[33] J. Wayman, "Biometrics in Identity Management Systems", *IEEE Security and Privacy*, Vol. 8, pp. 30-37, 2008.

[34] J. Vossaert, B. Lapon, De Decker and V. Naessens, "User-Centric Identity Management using Trusted Modules", *Mathematical and Computer Modelling*, Vol. 57, No. 57, pp. 1592-1605, 2013.

[35] F.G. Marmol, J. Girao and G.M. Perez, "TRIMS, A Privacy-Aware Trust and Reputation Model for Identity", *Computer Networks*, Vol. 54, pp. 2899-2912, 2010.

[36] D. Rutitis, A. Batragab and D. Ski, "Evaluation of the Conceptual Model for Corporate Identity Management in Health Care", *Proceedings of 19th International Scientific Conference; Economics and Management*, pp. 23-27, 2014.

[37] D. Rutitis and A. Batraga, "The Conceptual Framework of Corporate Identity in Health Care Industry", *Proceedings of International Symposium of Economic Crisis: Time for a Paradigm Shift-Towards A Systems Approach*, pp. 1-6, 2013.

[38] J. Wener, C.M. Westphall and C.B. Westphall, "Cloud Identity Management: A Survey on Privacy Strategies", *Computer Networks*, Vol. 122, pp. 29-42, 2017.

[39] R. Jamieson, P.L. Wee Land, D. Winchester, G. Stephens, A. Steel, A. Maurushat and R. Sarre, "Addressing Identity Crime in Crime Management Information System: Definitions, Classification, and Empirics", *Computer Law and Security Review*, Vol. 28, No. 4, pp. 381-395, 2012.

[40] H.R. Khedmatgozar and M.A. Hafezi, "The Role of Digital Identifier System in the Digital Objects", *International Journal of Information Management*, Vol. 37, pp. 162-165, 2017.

[41] P. Vartiainen, "On the Principles of Comparative Evaluation", *Evaluation*, Vol. 3, No. 8, pp. 359-371, 2002.

[42] T. J. Smedinghoff, "Solving the Legal Challenges of Trustworthy Online Identity", *Computer Law and Security Review*, Vol. 28, No. 5, pp. 532-541, 2012.

[43] R.M. Davison and C.X. Ou, "Digital Work in Digitally Challenged Organization", *Information and Management*, Vol. 54, No. 1, pp. 129-137, 2017.

[44] C. Sullivan, "Protecting Digital Identity in the Cloud: Regulating Cross Boarder Data Disclosure", *Computer Law and Security Review*, Vol. 32, No. 2, pp. 137-152, 2014.

[45] S. Aleem, L.F. Capretz and F. Ahmed, "A Digital Game Maturity Model (DGMM)", *Entertainment Computing*, Vol. 17, pp. 55-73, 2016.

[46] C. Sullivan, "Digital Identity-The Legal Person?", *Computer Law and Security Review*, Vol. 25, No. 4, pp. 227-236, 2009.

[47] L.E. Wang and L. Xianxian, "A Graph-based Multifold Model for Anonymizing Data with Attributes of Multiple Types", *Computers and Security*, Vol. 72, pp. 122-135, 2018.

[48] Y. Xu, K. Wang, A. Fu and P. Yu, "Anonymizing Transaction Databases for Publication", *Proceedings of 14th ACM International Conference on Knowledge Discovery and Data Mining*, pp. 1219-1227, 2008.

[49] G. Poulis, G. Loukides, A. Gkoulalas Divanis and S. Skiadopoulos, "Anonymizing Data with Relational and Transaction Attributes", *Proceedings of International Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 353-369, 2013.

[50] T. Takahashi, K. Sobataka, T. Takenouchi, Y. Toyoda, T. Mori and T. Kohro, "Top-Down Item Set Recording for Releasing Privacy Complex Data", *Proceedings of International Conference on Privacy, Security and Trust*, pp. 1-5, 2013.

[51] E. Torroglosa, J. Ortiz and A. Skarmeta, "Matching-Federation-Identities, the-eduGAIN and STORK Approach", *Future Generation Computer System*, Vol. 80, pp. 126-138, 2018.

[52] I. Gomaa, A.M. Said, E. Abd-Elrahman, A. Hamdy and E.M. Saad, "Performance Evaluation of Virtual Identity Approaches for Anonymous Communication in Distributed Environments", *Procedia Computer Science*, Vol. 109, pp. 710-717, 2017.

[53] J. Rell, "Mobile Cloud Security: Attribute-Based Access Control", *Mobile Cloud Computing*, Vol. 12, pp. 181-211, 2018.