IDRBT
Explore, Enable, Excel

Staff Paper Series

Vol. 5 | No. 1 | August 2019

FinTech

Data-oriented and Machine Learning Technologies in FinTech
- Dr. Mridula Verma, Assistant Professor 01

Recent Advances in Blockchain Technology
- Dr. N. V. Narendra Kumar, Assistant Professor 30

Emerging Security Solutions
- Dr. V. Radha, Associate Professor 55

Financial Inclusion: Emerging Role of FinTech
- Dr. M. V. N. K. Prasad, Associate Professor 85

Institute for Development and Research in Banking Technology
(Established by Reserve Bank of India)

Foreword

FinTechs and an Evolving Ecosystem

When FinTechs were emerging a few years back, there was a concern amongst banks about competition from them. The concern was, however, short-lived. Banks and FinTechs have started collaborating closely, paving the way for innovative banking products and services. An entire ecosystem led by regulators as well as both central and stage governments has been evolving. The ecosystem comprises academic institutions, incubators, accelerators, start-ups, major IT companies and funding institutions.

IDRBT, an institute working closely with academic institutions, banks and IT companies has taken a few initiatives to help a closer coordination among all of them. FinTech Forum is a platform created by IDRBT for enrolment of FinTechs willing to work closely with banks in the areas of interest and concern to banks. The forum aims to achieve coordination in different ways – regular meetings, an exchange for sharing of information by FinTechs with others, portal for innovative ideas of FinTechs and pain points of banks, and a test bed for FinTechs to test new ideas in the form of PoCs, prototypes, simulated exercises and test runs with the support of banks. We are confident the forum would over a period of time position itself as a repository of FinTech innovations useful for banks.

We observe that some of the major areas in which FinTechs have been working include payment systems, mobile banking, analytics, artificial intelligence, cyber security, customer interface, risk management and blockchain technology. Our researchers are studying these areas as part of their respective research centre activities. In order to put together the research studies undertaken by them in the areas relevant to FinTechs, we have chosen the theme for the current IDRBT Staff Paper Series as FinTech.

There are four articles in the present volume, contributed by Dr. Mridula Verma, Dr. N.V. Narendra Kumar, Dr. V. Radha and Dr. M.V.N.K. Prasad on Artificial Intelligence, Blockchain, Security and Financial Inclusion, respectively. The articles present the present status of academic work in the concerned areas along with opportunities for FinTechs to exploit them.

We trust that this Staff Paper Series will benefit the FinTechs in designing and deploying the appropriate solutions useful to banks and their customers.

Date: August 2019
Place: Hyderabad

Dr. A. S. Ramasastri
Director, IDRBT

Data-oriented and Machine Learning Technologies in FinTech

– **Dr. Mridula Verma,**
Assistant Professor, IDRBT

Abstract

With the continuous breakthroughs and growth in various areas of the financial sector, FinTech has achieved a great deal of attention. Banks and financial institutions are realizing the value-addition in adopting / adapting the fintech innovations for mutual and customer benefits. However, plenty of challenges exist – due to the intersection of two different fields, complex integrated systems, and expanded expectations. Thus, having the latest knowledge of FinTech is imperative for academicians as well as banking professionals and practitioners. In this article, we present a detailed discussion on few of the advanced data-oriented and machine learning based technologies.

1. Introduction

A relatively new industry in India, the Financial Technology (FinTech) industry primarily deals with a wide range of applications of information technology in the finance and banking sector. Few areas where new research ideas are successfully being applied to improve the quality of banking services include mobile networks [1-3], trust management [4-6], cloud computing [7-9], big data [10-12], image processing [13-15], and data analytics [16-18, 32]. Due to various developments in technologies, increasing expectations of business innovations, the requirement for affordable solutions, and growing customer demand – FinTechs are seeing investments from global financial firms [19]. The growing trend of investment in FinTech in the recent five years is depicted in **Figure 1**. **Figure 1(a)** shows the global trend and the trend of investment in India is shown in **Figure 1(b)**.

Fintech is accelerating the pace of change and reshaping the banking services in India radically. Banks and financial institutions are realizing the value-addition in adopting and/or adapting the fintech innovations for mutual and customer benefits. These collaborations present a great opportunity for banks and financial institutions in India in realizing innovative products and technologies. The RBI has also set

up a working group on FinTech and Digital Banking. According to a recent report [20] of the working group, FinTech innovations has been categorised into five major groups, namely: (i) Payments, Clearing & Settlement; (ii) Deposits, Lending & Capital Raising; (iii) Market Provisioning; (iv) Investment Management; and (v) Data Analytics & Risk Management, as shown in **Figure 2**. Out of these five major technical dimensions, this article covers some of the crucial aspects and interesting use-cases of data-oriented and machine learning based technologies. It is important to note that in addition to these five cubes of FinTech innovations, one more dimension that passes through all these dimensions is privacy and security, which plays an important role in adoption of FinTech by practitioners.

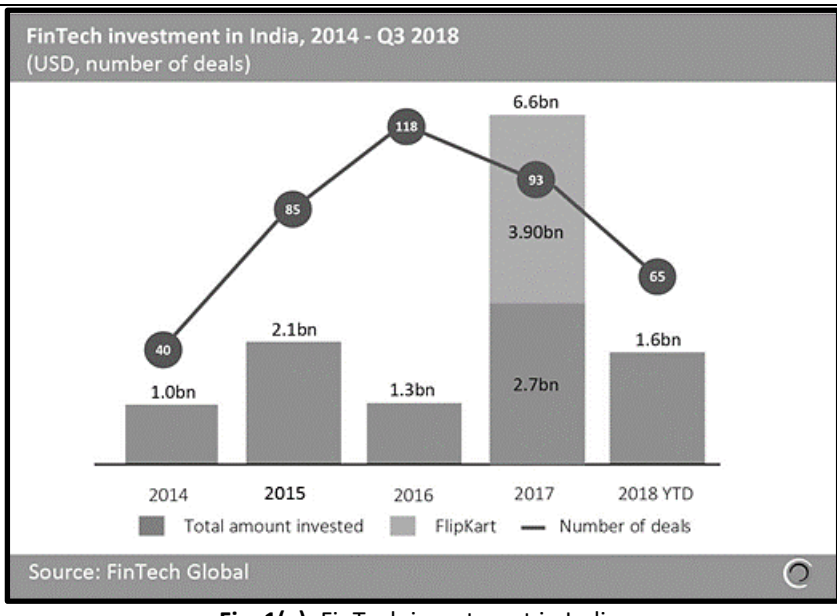


Fig. 1(a): FinTech investment in India

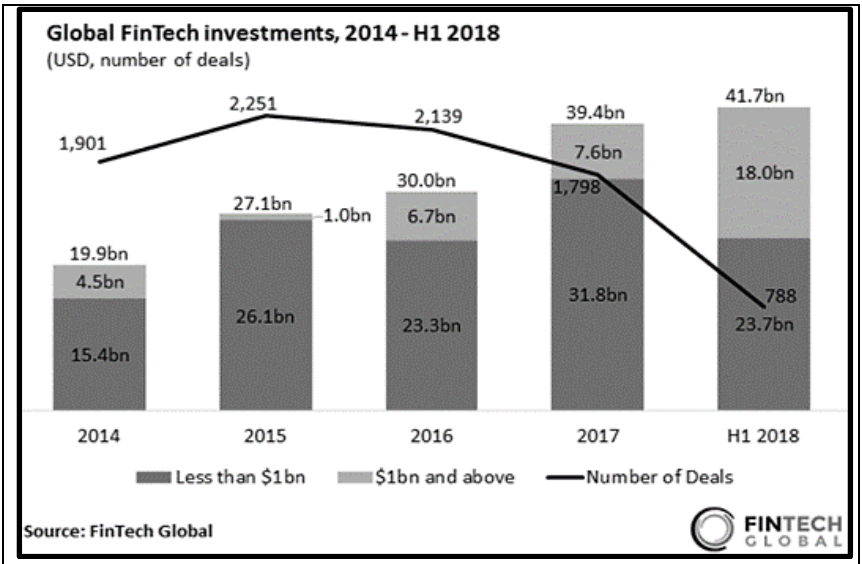


Fig. 1(b): Global FinTech Investments

Fig.1: The trend of investment in FinTech

Payments, Clearing & Settlement	Deposits, Lending & capital raising	Market provisioning	Investment management	Data Analytics & Risk Management
Mobile and web-based payments Digital currencies Distributed ledger	Crowd-funding Peer to peer lending Digital currencies Distributed Ledger	Smart contracts Cloud computing e-Aggregators	Robo advice Smart contracts e-Trading	Big data Artificial Intelligence & Robotics

Fig. 2: Categorization of major FinTech Innovations

According to FinTech Outlook report for 2018 [55], seven key technologies are driving current FinTechs: Cloud, IoT, Cyber Security, Biometrics, AI, Mobility, and Big Data and Analytics. Figure 3 shows a few of the popular and active areas where FinTech opportunities are available in India. All data-oriented operations for the purpose of financial and banking services come under the class of data mining and data analytics technologies. In the area of data mining and analysis, the main objective is to extract useful knowledge from the raw data and then use this knowledge for value creations. Data-oriented techniques can be classified into two categories, namely, (i) big data processing, and (ii) data analytics and mining. Techniques that

come under the first category are about the processing or modeling of large volume of data that impact quantitative finance. The techniques that belong to the latter class are more centered on the goal of retrieving more information from the financial and banking data in large volumes.

Artificial Intelligence (AI) and Machine Learning (ML) techniques are emerging as promising solution providers, which have the potential to improve customer service and in-house processing in financial institutions and banks. In the view of growing success-rate of research and development of AI and ML based tools and technologies in various application domains, many FinTechs are now turning to develop automated self-learning algorithms for improved and efficient financial solutions. In contrast to early days, when banking and financial professionals employed analytics at different stages of decision-making, now banks are looking forward to having smart solutions to achieve improved efficiency of data analytics algorithms, better customer credit analysis, accurate decision-making, boosted security and improved customer experience. With the help of different analytical and intelligent models, FinTechs are now widely adapting self-learning methodologies. In India, there is a large scope for investment in these areas to automate the workflow of the system, apply smart analytical solutions, and gain efficient customer service.

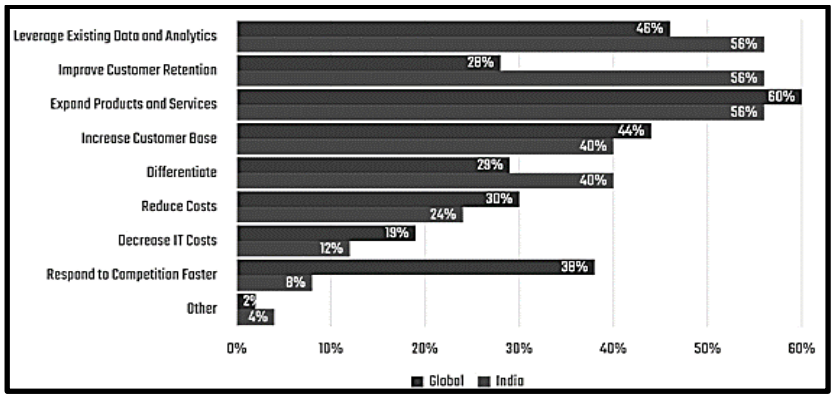


Fig. 3: Most promising FinTech Opportunities in India (Source: FinTech Trend Report India 2017, PwC)

A few use-cases of ML in banking sector includes voice and image recognition, improved search and matching capabilities, improved logistics, supply chain management, and personalised marketing. In the past five years, a good number of FinTech companies are engaged in research, development and innovations in technology areas like natural language processing, deep learning, and social network analysis using graph-based data mining algorithms, to create solutions that generate greater value. One interesting example is using ML and NLP to analyse online financial textual content to predict the stock market movements. Due to a number of microblog websites such as Twitter and Stocktwits, and online news websites such as Reuters, etc., abundant amount of textual data is available. A number of research labs are currently working to analyse the correlation between the sentiments extracted from financial news and microblogs on the stock price movement.

In this article, we will first discuss data analytics and big data processing tools and techniques in more detail in the succeeding section. In addition, we will also discuss few applications, such as intelligent micro conversations, robo-advisors, credit risk assessment,

and biometrics-based security systems in BFSI, where AI and ML based technologies are being adopted by Indian FinTech companies.

2. Data Analytics and Big Data Processing

With exponential technological growth, an enormous amount of data is generated on a daily basis. The primary objective of financial data analytics and mining is to extract knowledge from this large volume of available data, profitably. In [21], the author discussed various aspects of 'large-volume, complex, growing data sets' with multiple autonomous sources. In order to explain the characteristics of the big data, the author describes the HACE theorem as follows:

HACE Theorem*: Big Data starts with large-volume, heterogeneous, autonomous sources with distributed and decentralized control, and seeks to explore complex and evolving relationships among data.*

Using the allegory of a blind man and a giant elephant, the author presented the main issue of localized view of the whole data, as shown in **Figure 4**. This problem is faced by the big data practitioners, since a localized view only provides a partial analytical result, which might be different from the nature of actual inference that could be drawn if the data is processed on the whole.

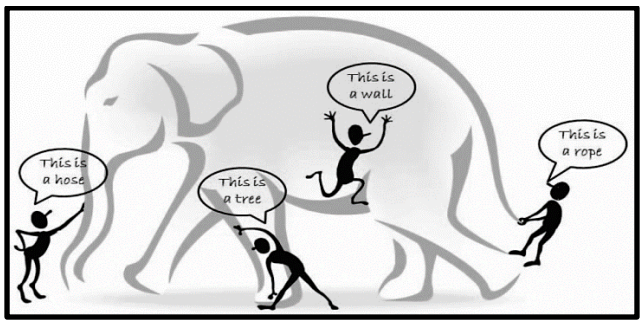


Fig. 4: The limited view of each blind man leads to a biased conclusion [21].

In addition to the above theorem, 5V characteristics (Volume, Velocity, Variety, Veracity, and Value) popularly known in the

literature are described in **Figure 5**. In the view of the wide scope of big data analytics and mining, practitioners are looking forward to the usage of large data centers. However, this approach significantly affects the centralized data computing system, due to the conflicts between energy efficiency and computational performance [44]. Some recent research addresses this issue with contemporary technologies such as green computing and parallel architecture. For instance, in the field of green computing, the Phase-change Memory (PCM) is used as a promising scalable technique, which can replace the non-scalable dynamic random access memory. In [45], the authors proposed a genetic algorithm to improve the performance of the PCM. Similarly, in [46], a heuristic algorithm is proposed for allocating the data in the heterogeneous memory. Another technology in the area of parallel architecture is the Field-Programmable Gate Array (FPGA) based architecture proposed in [47]. The authors showed the performance of the proposed architecture on the problem of detection of human behaviour.

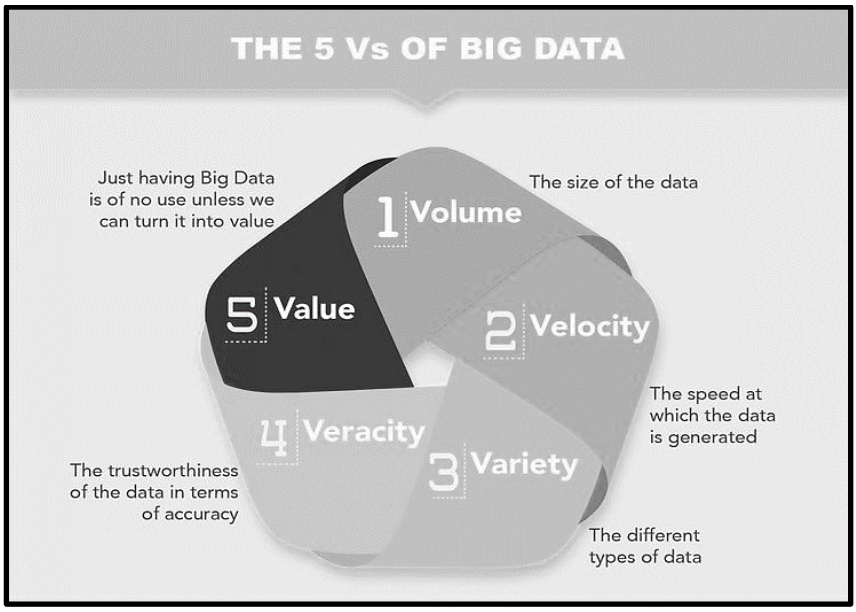


Fig. 5: The 5Vs of Big Data [54]

Technologies based on the statistics of the data for modeling the business process and other quantitative analysis on the financial data are also getting updated to support big data framework. The technological trend in FinTech is migrating from elementary statistics and structured data processing to the complex computations and semi-structured or unstructured data handling. There are a number of dimensions available in the big data computation such as distributed computing, blockchain, unstructured data analysis and many more, where the contemporary technologies are being developed [41-43]. Various other technologies that handle the computations on big data environment include Hadoop MapReduce [48][52], Spark [50], Cloud Computing [49][53], and Smart Grid [51], etc.

In the field of banking and finance, a number of applications use data analytics and data mining as the base technology. A few examples include forecasting stock market, financial risk prediction, credit rating, money laundering, bank customer profiling. With the growth in the volume of available data, the research and development in the field of big data analytics is also increasing. Few of these techniques are discussed in [22] and [23]. For the quantitative data analysis on the financial datasets, in [23], the authors discussed different mathematical and statistical tools and computing technologies to perform the process of cleaning, aggregation and the modeling of the data. The paper also discusses the development of the quantitative analysis in finance data. Recently, in order to perform the quantitative analytics to big financial datasets on the cloud, in [24], the authors proposed a new technology called QuantCloud. The QuantCloud infrastructure employs a large-scale SSD-backed data storage and using a number of parallel processing algorithms, it facilitates data-on-demand service.

Another important technological requirement in the banking and finance sector is in providing the customized or personalized services to clients, such as private financial advice and management. For such type of applications, it is required to use the contextual information

from heterogeneous metadata in order to create personalized recommendations. In order to gather such metadata, the general source of information is the online social networks, web queries and public sentiment. These sources provide useful/distinctive information in terms of features based on which the data analysis/data mining models are generated. Of late, a new technique has been proposed in [25], which considers the trustworthiness of user ratings, spatial-temporal features, and reviews sentimental features to evaluate the service model. Another technique is given in [26], which discusses the task of recommending a personalized travel sequence based on the contextual information extracted from tags, geolocation and time of photographs shared in online travelogues and social communities such as Flickr.

As discussed earlier, in the banking and financial sector, the fields of artificial intelligence and data analytics centre around the exploration and investigation of data to extract useful patterns and knowledge, which can be applied to improve financial services. One example of these services is fraud detection, which is mostly used by auditors for detecting the fraudulent transaction from millions of transactions [27]. In addition to this work, many data mining and machine learning based-approaches are applied for this task [28-30]. Such patterns also describe different types of relationships between financial entities, which lead to major financial incidents. Ontology-based approaches are one instance of such semantic solutions, which rank the web services in order to enhance the discoverability. One application of such an ontology-based multi-agent system is discussed in [31].

When a large amount of user data is publicly available, extracting the sensitive information is not difficult for an intruder, and thus preserving the privacy of the user becomes a critical issue. When a knowledge extraction technique is applied to user's data, there exists a high risk of discovery of sensitive information about that user. Hence, it is important to preserve the privacy of the user while extracting useful information from the data. Traditional methods in

this direction rely on the concepts of perturbation and/or randomization [33-36] and cryptographic tools [37-38]. More advanced frameworks are designed where the secure deep learning frameworks are being proposed [39] and the classification task is applied over the encrypted data [40]. Therefore, more privacy-preserving data analytics and machine learning algorithms and technologies are required in the coming future.

3. AI and ML Technologies used by Indian FinTechs

As discussed earlier, AI and ML based technologies are one of the key technologies, driving FinTechs to produce better and smarter customer services and financial solutions. In this section, a brief review of the technologies that are being used in four different trending applications is explained:

3.1. Intelligent Micro-conversations: Chatbots

Chatbots are now popularly being used for understanding the customers better and gathering information about their requirements. These chatbots are virtual assistants that are in general an adaptive algorithm, which works as conversational partner to humans. In banking sector, bots, which are now termed as Bankbots, communicate with customers in their language, understand their specific requirements and provide financial services accordingly. The other uses and benefits of bankbots includes – customer engagement for the new account generation, 24/7 digital support, and cost saving. A Gartner report predicts that by 2020, chatbots would support around 85% of the overall client support services. **Figure 6** describes the required properties of a well-designed chatbot.

The technologies behind a textual chatbot include adaptive machine learning, natural language understanding and natural language processing techniques along with context-sensitive processing, analysis of sentiments involved in the textual messages, and adaptive learning capabilities. The sources of the data for a chatbot program

not only includes structured historical transactional log data, and management information system data, but also the unstructured social media data and user's purchase behaviour data. With the advent of technologies, processing power, complex deep neural network learning architectures are being utilized to identify the associated sentiments in user's text messages and predict the relevant reply from the bank's end. A few of the popular chatbots are Watson Assistant, Bold360, Rulai, LivePerson, Inbenta, Ada, and Vergic.

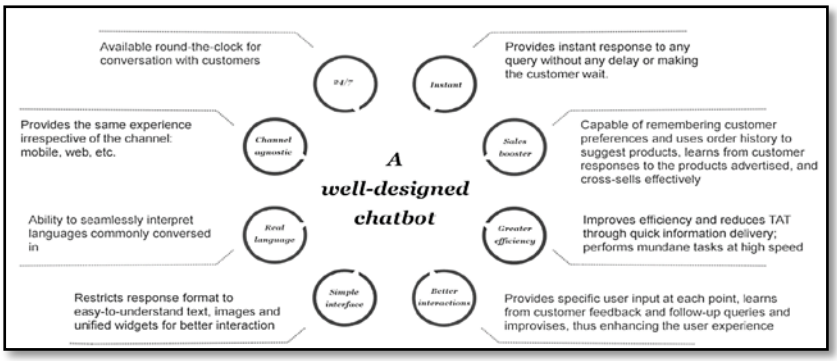


Fig. 6: Required properties of a Chatbot [56]

In the process of implementation of text-based chatbots, it is essential to apply different NLP-based pre-processing techniques to digest the data. In Python programming language, the Natural Language Toolkit (NLTK) and sklearn packages are very popular for the implementation of such pre-processing methods. Following are few pre-processing methods that can be applied to the textual data:

Method 1 – Cleaning:

- ***Tokenization:*** In this process, each paragraph is partitioned into smaller parts, in the form of sentences or words, called sentence tokens or word tokens respectively. **Figure 7** shows the process of both types of tokenization.

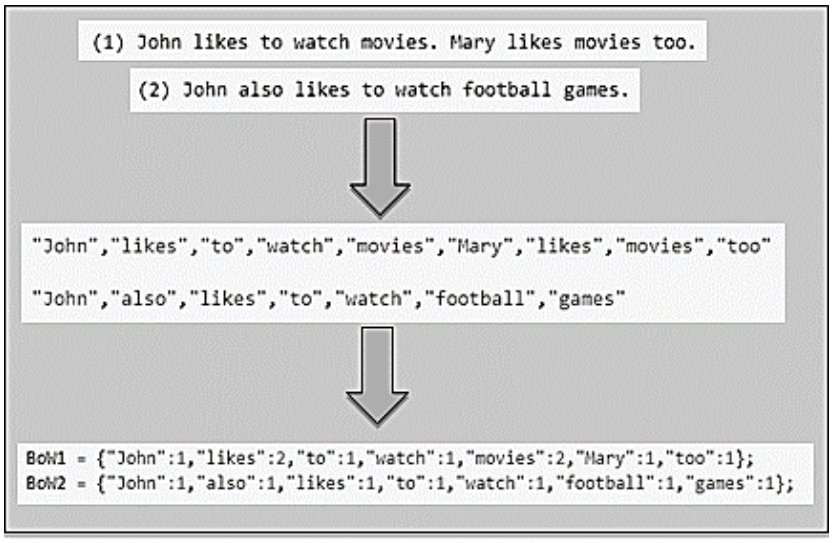


Fig. 7: Tokenization and Bag-of-Word representation

- ***Stemming:*** In this process, all the words are reduced to its root form (stem), by removing the affixes. This process is shown in **Figure 8(a)**. We can also consider this process as a part of textual normalization

- ***Lemmatization:*** Linked to the process of stemming, this process captures the dictionary form of the word. In **Figure 8(b)**, we have given an example of this process

Fig. 8(a): Stemming

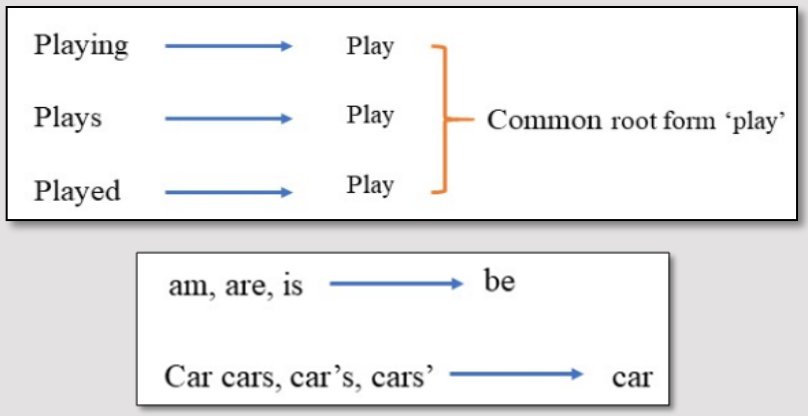


Fig. 8(b): Lemmatization

Fig. 8(c): Example sentence before and after Stemming and Lemmatization

Fig. 8: Stemming and Lemmatization [58]

- ***Removing Stopwords:*** Considered as noise, stopwords are those commonly used words, which do not add any meaning to a sentence, and occurs for just the grammatical precision. Few examples are "a", "an", "the", "is", "am", etc.

- ***Parts of speech tagging:*** In this process, based on the context, each word in the corpus is tagged (or linked or marked-up) to the part of speech to identify the group of grammar, i.e. Noun, Pronoun, Adjective, Verb, Adverbs, etc. **Figure 9** describes this process

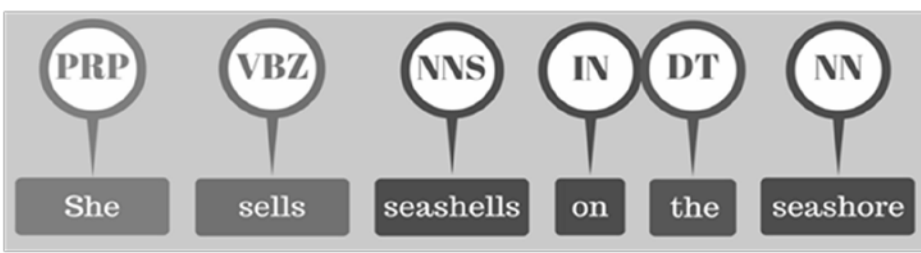


Fig. 9: PoS Tagging

Method 2 – Feature extraction involves extracting numerical features from the textual documents. Few of these features include:

- **TF-IDF:** TF stands for term frequency, which computes how frequently a term occurred in a document. For term "t", it is computed as follows,

$$\text{TF(t)} = \frac{Number\ of\ times\ 't'\ appears\ in\ a\ document}{total\ number\ of\ terms\ in\ the\ document}$$

IDF stands for Inverse Document Frequency, defined as $\log_e$ of ratio of number of total documents in the corpus and the number of documents containing the term "t",

$$\text{IDF(t)} = \log_e \frac{Total\ number\ of\ documents}{Number\ of\ documents\ with\ term\ 't'\ in\ it}$$

The TF-IDF measures the relative importance of a term in a corpus (list of documents), computed by TF(t) × IDF(t). The sklearn package of Python programming language is used to extract the TF-IDF of text

- **One hot encoding:** This encoding is used to represent the categorical variables as binary vectors. Example is shown in **Figure 10**. Each word is represented by a binary vector, with all zero values except for the index of the word

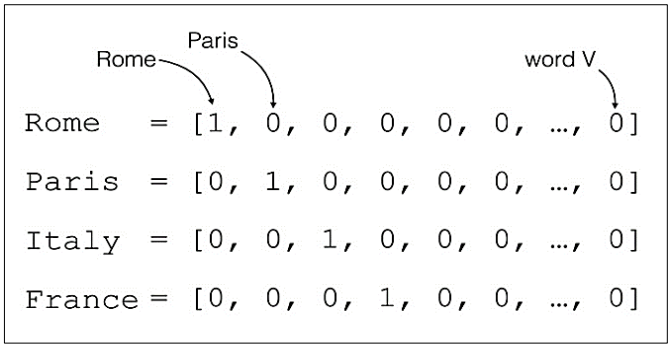


Fig. 10: One hot encoding [59]

- **Bag-of-Words Model:** BoW model is a type of representation, used to store textual documents. This model stores the number of times one word has occurred in a document, disregarding the meaning and order of the words as shown in **Figure 7**
- **Word2vec:** The term Word2vec is given to a set of models that is used to learn the vector representations of words, known as the "word embedding". After getting the vector representations, these vectors are sent as the input to different machine learning models. A large corpus of text is taken as input and output is generated as a vector space of a pre-specified number of dimensions, where each unique word in the corpus is assigned with a corresponding vector in the space. Word vectors are positioned in the vector space such that words that share common contexts in the corpus are located in close proximity to one another in the space [57]
- N-grams are defined as a continuous sequence of words, where 'n' specifies the length of the sequence. The n-gram tokenizer first breaks text down into words, whenever it encounters with one of the list of specified characters, then it emits n-grams of each word of the specified length.

After extracting the numerical features from the textual dataset, these features are used to train different machine learning classifiers. Few of these classification models are discussed in the IDRBT Staff Paper on Analytics [61]. Context sensitive processing in chats include correction of the spelling and expansion of the acronyms used in textual messages. Most of these methods use some standard dictionaries, which are designed based on the domain knowledge of such documents and different classifiers such as SVM, Logistic Regression, etc. Domain based dictionaries remove the ambiguities in the process of expansion of acronyms. One example of such ambiguity is in the expansion of the word "CMU", which can have different meanings such as, "Carnegie Mellon University", "Current Monetary

Unit", "Central Money-market Unit", "Crisis Management Unit" and many more.

One of the most important processes in chatbot programs is to analyze the semantics of the replies, which is performed by understanding the natural language. After getting the content, this process tries to capture the real meaning of the text. This step helps in identifying the relevance of the text as well as understanding the relationship between different concepts.

Another technology used in such textual analysis is the sentiment recognition process, which detects the orientation of the semantics in the textual replies. These methods are very important modules in the chatbot program, as these methods learn and provide complex hidden knowledge/structure/polarity in the text. Approaches for sentiment analysis include knowledge-based techniques, statistical methods and hybrid methods. Knowledge-based techniques use different classifiers to classify text based upon the presence of unambiguous impact words such as happy, sad, afraid, and bored. Statistical methods also leverage elements from machine learning such as latent semantic analysis, SVM, "bag-of-words", "Point wise Mutual Information", for identifying the semantic orientation. Currently, a number of deep learning frameworks are being used for this process. Hybrid approaches leverage both machine learning and elements from knowledge representation. Recursive neural networks are used in sentiment analysis to perform the best, specifically on review-type sentiment analysis with text ranging from one sentence to one paragraph.

Involving different IoT devices and video-and-speech processing technologies – which support recognition of activities, gestures, spoken words and the language used by the user – these chatbots are now reaching much advanced standards. Here the sentiments of the customers can be analyzed using the tone, facial expressions, and voice accent. There are however, numerous challenges in implementing an effective chatbot due to several limitations. These

include the limited dialog capability as well as the implementation for multi-lingual regions.

3.2. Robo-Advisor

Just like a human financial advisor, machine learning algorithms can now generate automated financial advices and online investment management tools for the customers. These tools require little to no interventions from humans and with the help of a short survey about the customer's financial position and goals, they provide algorithm based financial planning services. These machine learning algorithms and deep learning frameworks are able to identify complex hidden structures that which may go undetected by human eyes. These latent patterns are helpful in selection of the investments and diversified portfolio building, automatically. With respect to the target allocation, to form an optimal portfolio, the robo-advisor software can dynamically adjust the investments. Tax-loss harvesting is one of the best features provided by robo-advisors.

ML and AI can provide support not only to the process of portfolio management, but also enhance the user's experiences by providing text and voice based technologies. In this area, deep learning frameworks, such as Long Short-Term Memories Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) are giving good results. Recent research directions are applying advanced ML concepts, such as Kernel-based learning, Extreme ML [60]. In addition, researchers are also exploring the directions to employ personalized recommendations and insights from the social behavior of the user, which can be extracted by analyzing the social network of the user.

3.3. Credit Risk Assessment

In this digital lending space, where people with no credit history, online, apply for a loan, assessing their credibility is becoming a major challenge. Solving this problem using AI and ML techniques is currently trending in FinTech sector. For such type of analysis, credit

score associated with a loan application is computed for which FinTechs use data, based on the digital footprints, social network, psychometric test scores, and other demographic-, social- and business- data of the applicant. AI application could also help suggest the perfect lenders to the suitable borrower.

The credit risk is defined as the risk associated with a loan application that may emerge if the applicant fails to pay the loan amount. The calculation of the credit risk for a borrower is very crucial for a bank before taking the decision to accept or reject a loan application. The measure, which is used for such creditworthiness of one borrower, is called as credit score.

In current scenario, with the advent of AI and ML based technologies and large amount of available data, the application of credit risk modelling has become a useful tool, with help of banks to alleviate the overall risks. The data that is used in this task includes – the current balance in the applicant's account, loan balance, age, salary, profession, assets, payment history, number of dependents, etc. At the same time, ML based approaches available to handle this problem of credit risk modelling includes – random forest, gradient boosting and deep learning frameworks such as Convolutional Neural Networks.

The random forest model to predict the credit score of an applicant is made up of ensembling of multiple decision tree classifiers, each trained with randomly selected subspaces of the whole data. The accuracies of individual decision tree classification models shape up the final accuracy we obtain from a random forest built on those individual trees jointly. In **Figure 11**, the prediction by one random forest is shown on a two dimensional sample dataset, having two different labels/classes, which can be considered as the accepted and rejected loan applications.

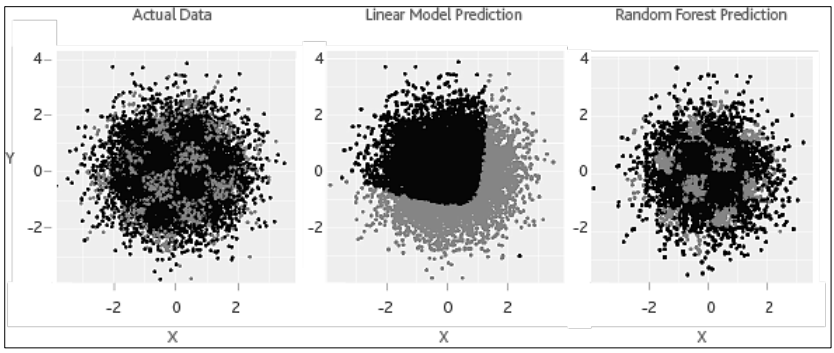


Fig. 11: Prediction performance of Random Forest [62]

Gradient boosting is another ML technique that is used for regression and classification. Here again, the few weakly predicting classifiers, typically decision trees are ensembled. This method follows the traditional loss or cost minimization framework of machine learning. The cost function in machine learning is a function, which represents the difference between the actual value and the predicted value. One limitation of this method, nonetheless, is found to be the overfitted models. Overfitting is the concept where the trained model works very well, i.e. provides lesser error with the training data, however, with the test data, or in real-run, the accuracy of the model is reduced.

XGBoosting, is a distributed gradient boosting library, which is used by many practitioners. It provides framework for the parallel running of multiple trees that runs on distributed environment. Apart from the traditional gradient boosting technique, it also includes stochastic as well as the regularized gradient boosting. The stochastic gradient boosting method uses sub-sampling of the whole dataset that makes the process significantly faster, whereas, regularized gradient boosting reduces the overfitting of the traditional method.

Deep learning frameworks, such as deep belief networks, recurrent neural network, and long short-term memory network are also being applied to solve the problem. These models are also good in identifying the hidden structure in the alternative source of data, such as social profile data, rental data, social behavioral data, etc.

3.4. Biometrics-based Security Systems

The three major categorizations of the user authentication approaches are knowledge based approaches (with the help of pin and password), possession based (smart card, QR code), and biometrics-based approaches. Biometrics-based security system is currently one of the root security systems in most of the banks to authenticate and identify person in a trustworthy manner. Fingerprint, iris, face, voice-pitch, palmprint, are few examples of the physical biometrics, that are being utilized for providing security to customers at different platforms. Few instances of usages are: a) DCB bank using fingerprints to withdraw money from ATM; b) Federal bank using zero-balance selfie account; and c) ICICI's using of voice recognition technology based authentication. Biometrics based security systems are getting more popular for providing a baseline to the customer's security, as it is quite hard for the customers to remember the PIN/Passwords for different accounts. In addition to the physiological-biometrics, the behavioral-biometrics, such as keystroke dynamics, gait, clickstream, are also being used for different purposes, such as customer authentication.

AI and Machine learning based algorithms are being used at different stages of authorization systems. Few algorithms are devising best set of distinguishing features at the feature extraction phase, whereas other algorithms are used to identify hidden complex patterns in datasets, which can be used for distinguishing between genuine customers and fraudsters. Recently, using few advanced CNN models, better features were extracted [64], [65]. In [63], which are used to learn the reduce size of physiological biometric templates. Similarly, fusion based feature extraction [66] [67] is also an interesting direction of research, where multiple views, or multiple modalities or the complimentary information of biometric data are combined together to design a better set of features.

With the advent of internet technology, the availability of physiological data of a user is lesser than the behavioral properties. In

Figure 12, few examples of biometrics data are shown, which can be easily extracted from a mobile phone. Refer [71] for more details.

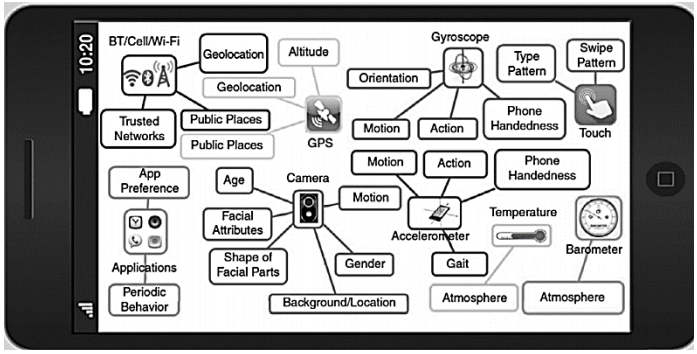


Fig. 12: Example of Behavioral Biometrics [73]

In recent research, combining the data extracted from multiple sources, or combining the multiple modality data together has been proved to be useful in designing better security. For example, the features extracted from the clickstream for a user can be combined with the features extracted from the keystroke dynamics, such that the resultant feature set will contain the properties of both the modals. Similarly, learning can be applied in multimodal fashion. For such type of fusion, according to recent literature, deep learning approaches are outperforming other methods [66][72]. For a detailed survey on the deep learning methods in this area, do refer [68].

Although various ML and AI based technologies are available, there is a dire need of techniques that preserves the privacy of the user. A few privacy-preserving machine-learning approaches were proposed, where the learning is performed on the encrypted data [69][70]. Such type of approaches can also be applied to make a secure analytics system. In addition, although deep learning frameworks are currently state-of-the-art approaches, there are few limitations, such as deep learning architecture for cross-domain data, or a large amount of dynamic data.

Conclusion

In this article, we presented a detailed discussion on few of the advanced data oriented and machine learning based technologies. We discussed few important concepts of big data analytics as well as a set of popular use-cases where the AI and machine learning based techniques are being applied to produce better as well as smarter customer services and financial solutions. We also discussed a few open problems, which might help in designing improved technologies.

References

1. Qiu, M., Gai, K., Thuraisingham, B., Tao, L., & Zhao, H. Proactive user-centric secure data scheme using attribute-based semantic access controls for mobile clouds in the financial industry. Future Generation Computer Systems. 80, 421-429, 2018
2. Wang, Y., Wen, S., Xiang, Y., & Zhou, W. Modeling the propagation of worms in networks: A survey. IEEE Communications Surveys & Tutorials, **16**(2), 942-960, 2014
3. Panwar, N., Sharma, S., & Singh, A. K. A survey on 5G: The next generation of mobile communication. Physical Communication, 18, 64-84, 2016
4. Zhang, Q., Yang, L. T., & Chen, Z. Privacy-preserving deep computation model on the cloud for big data feature learning. IEEE Transactions on Computers, **65**(5), 1351-1362, 2016
5. Li, P., Li, J., Huang, Z., Li, T., Gao, C. Z., Yiu, S. M., & Chen, K. Multi-key privacy-preserving deep learning in cloud computing. Future Generation Computer Systems. 74, 76-85, 2017
6. Anjum, A., Sporny, M., & Sill, A. Blockchain standards for compliance and trust. IEEE Cloud Computing. **4**(4), 84-90, 2017
7. Gai, K., Qiu, L., Chen, M., Zhao, H., & Qiu, M. SA-EAST: security-aware efficient data transmission for ITS in mobile heterogeneous cloud computing. ACM Transactions on Embedded Computing Systems (TECS), **16**(2), 60, 2017
8. Li, Y., Gai, K., Qiu, L., Qiu, M., & Zhao, H. Intelligent cryptography approach for secure distributed big data storage in cloud computing. Information Sciences. 387, 103-115, 2017

9. Gai, K., Qiu, M., & Zhao, H. Energy-aware task assignment for mobile cyber-enabled applications in heterogeneous cloud computing. Journal of Parallel & Distributed Computing. 111, 126-135, 2018
10. Lei, Y., Jia, F., Lin, J., Xing, S., & Ding, S. X. An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data. IEEE Transactions on Industrial Electronics. **63**(5), 3137-3147, 2016
11. Yin, S., & Kaynak, O. Big data for modern industry: challenges and trends [point of view]. Proceedings of the IEEE, **103**(2), 143-146, 2015
12. Genuer, R., Poggi, J. M., Tuleau-Malot, C., & Villa-Vialaneix, N. Random forests for big data. Big Data Research. 9, 28-46, 2017
13. Castiglione, A., De Santis, A., & Soriente, C. Taking advantages of a disadvantage: Digital forensics and steganography using document metadata. Journal of Systems and Software. **80**(5), 750-764, 2007
14. Raghavan, S. Digital forensic research: current state of the art. CSI Transactions on ICT. **1**(1), 91-114, 2013
15. Guo, Q., Zhang, C., Zhang, Y., & Liu, H. An efficient SVD-based method for image denoising. IEEE Transactions on Circuits and Systems for Video Technology. **26**(5), 868-880, 2016
16. Chen, H., Chiang, R. H., & Storey, V. C. Business intelligence and analytics: from big data to big impact. MIS quarterly, 1165-1188, 2012
17. Vasarhelyi, M. A., Kogan, A., & Tuttle, B. M. Big Data in accounting: An overview. Accounting Horizons. **29**(2), 381-396, 2015
18. Lv, Y., Duan, Y., Kang, W., Li, Z., & Wang, F. Y. Traffic flow prediction with big data: A deep learning approach. IEEE Trans. Intelligent Transportation Systems, **16**(2), 865-873, 2015
19. Wigglesworth, R. Fintech: Search for a super-algo, URL = https://www.ft.com/content/5eb91614-bee5-11e5-846f-79b0e3d20eaf, January 2016
20. https://rbidocs.rbi.org.in/rdocs/PublicationReport/Pdfs/WGFR68AA1890D7334D8F8F72CC2399A27F4A.PDF
21. Wu, X., Zhu, X., Wu, G. Q., & Ding, W. Data mining with big data. IEEE transactions on knowledge and data engineering. **26**(1), 97-107, 2014

22. Sun, Y., Shi, Y., & Zhang, Z. Finance Big Data: Management, Analysis, and Applications, 2019
23. Shi, X., Zhang, P., & Khan, S. U. Quantitative Data Analysis in Finance. In Handbook of Big Data Technologies. pp. 719-753, Springer, Cham, 2017
24. Zhang, P., Yu, K., Jessica, J. Y., & Khan, S. U. QuantCloud: big data infrastructure for quantitative finance on the cloud. IEEE Transactions on Big Data. **4**(3), 368-380, 2018
25. Zhao, G., Qian, X., Lei, X., & Mei, T. Service quality evaluation by exploring social users' contextual information. IEEE Transactions on Knowledge and Data Engineering. **28**(12), 3382-3394, 2016
26. Jiang, S., Qian, X., Mei, T., & Fu, Y. Personalized travel sequence recommendation on multi-source big social media. IEEE Transactions on Big Data. **2**(1), 43-56, 2016
27. Cao, M., Chychyla, R., Stewart, T., 2015. Big data analytics in financial statement audits. Account. Horiz. **29**(2), 423–429
28. Dal Pozzolo, A., Boracchi, G., Caelen, O., Alippi, C., & Bontempi, G. Credit card fraud detection: A realistic modeling and a novel learning strategy. IEEE transactions on neural networks and learning systems. **29**(8), 3784-3797, 2018
29. Carcillo, F., Dal Pozzolo, A., Le Borgne, Y. A., Caelen, O., Mazzer, Y., & Bontempi, G. SCARFF: A scalable framework for streaming credit card fraud detection with Spark. Information fusion. 41, 182-194, 2018
30. Min, W., Tang, Z., Zhu, M., Dai, Y., Wei, Y., & Zhang, R. Behavior Language Processing with Graph-based Feature Generation for Fraud Detection in Online Lending. Workshop on MIS2: Misinformation and Misbehavior Mining on the Web, USA, 2018
31. Ying, W., Ray, P., & Lewis, L. A methodology for creating ontology-based multi-agent systems with an experiment in financial application development. 46th Hawaii International Conference on System Sciences (HICSS). pp. 3397-3406, January 2013
32. Chen, C. P., & Zhang, C. Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences*, *275*, 314-347
33. Xu, K., Yue, H., Guo, L., Guo, Y., & Fang, Y. Privacy-preserving machine learning algorithms for big data systems. IEEE 35th

International Conference on Distributed Computing Systems (ICDCS). pp. 318-327, June 2015

34. Chaudhuri, K., Monteleoni, C., & Sarwate, A. D. Differentially private empirical risk minimization. Journal of Machine Learning Research. **12**(Mar), 1069-1109, 2011
35. Chaudhuri, K., & Monteleoni, C. Privacy-preserving logistic regression. Advances in Neural Information Processing Systems. pp. 289-296, 2009
36. Fong, P. K., & Weber-Jahnke, J. H. Privacy-preserving decision tree learning using unrealized data sets. IEEE Transactions on Knowledge and Data Engineering, **24**(2), 353-364, 2012
37. Laur, S., Lipmaa, H., & Mielikäinen, T. Cryptographically private support vector machines. 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, pp. 618-624, August 2006
38. Clifton, C., Kantarcioglu, M., Vaidya, J., Lin, X., & Zhu, M. Y. Tools for privacy-preserving distributed data mining. ACM Sigkdd Explorations Newsletter, **4**(2), 28-34, 2002
39. Shen, S., Tople, S., & Saxena, P. Auror: defending against poisoning attacks in collaborative deep learning systems. In Proceedings of the 32nd Annual Conference on Computer Security Applications, pp. 508-519. ACM, December 2016
40. Bost, R., Popa, R. A., Tu, S., & Goldwasser, S. Machine learning classification over encrypted data. NDSS. February 2015
41. Wang, H., Xu, Z., Fujita, H., & Liu, S. Towards felicitous decision making: An overview on challenges and trends of Big Data. Information Sciences. 367, 747-765, 2016
42. Casado, R., & Younas, M.. Emerging trends and technologies in big data processing. Concurrency and Computation: Practice and Experience, **27**(8), 2078-2091, 2015
43. Fang, B., & Zhang, P. Big data in finance. Big data concepts, theories, and applications, pp. 391-412. Springer, Cham, 2016
44. Qiu, M., Ming, Z., Li, J., Gai, K., & Zong, Z. Phase-change memory optimization for green cloud with genetic algorithm. IEEE Transactions on Computers, **64**(12), 3528-3540, 2015
45. Gai, K., Qiu, M., Zhao, H., & Qiu, L. Smart energy-aware data allocation for heterogeneous memory. In High Performance Computing and Communications; IEEE 14th International Conference on Smart City; IEEE 2nd International Conference

on Data Science and Systems (HPCC/SmartCity/DSS), IEEE 18th International Conference pp. 136-143, December 2016

46. Li, Y., Gai, K., Qiu, M., Dai, W., & Liu, M. Adaptive human detection approach using FPGA-based parallel architecture in reconfigurable hardware. Concurrency and Computation: Practice and Experience. **29**(14), e3923, 2017
47. Dittrich, J., & Quiané-Ruiz, J. A. Efficient big data processing in Hadoop MapReduce. Proceedings of the VLDB Endowment. **5**(12), 2014-2015
48. Ji, C., Li, Y., Qiu, W., Awada, U., & Li, K. Big data processing in cloud computing environments. In Pervasive Systems, Algorithms and Networks (ISPAN), IEEE 12th International Symposium, pp. 17-23, December 2012
49. Zaharia, M., Xin, R. S., Wendell, P., Das, T., Armbrust, M., Dave, A., & Ghodsi, A. (2016). Apache spark: a unified engine for big data processing. Communications of the ACM, **59**(11), 56-65
50. Song, Y., Zhou, G., & Zhu, Y. Present status and challenges of big data processing in smart grid. Power System Technology, **37**(4), 927-935, 2013
51. Chen, Y., Alspaugh, S., & Katz, R. Interactive analytical processing in big data systems: A cross-industry study of mapreduce workloads. Proceedings of the VLDB Endowment, **5**(12), 1802-1813, 2012
52. Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Khan, S. U. The rise of "big data" on cloud computing: Review and open research issues. Information systems, 47, 98-115, 2015
53. http://bigdata.black/featured/what-is-big-data/
54. http://www.opusconsulting.com/wp-content/uploads/fintech-outlook-2018.pdf
55. https://www.pwc.in/consulting/financial-services/fintech/fintech-insights/chatbot-the-intelligent-banking-assistant.html
56. Mikolov, T., Chen, K., Corrado, G., & Dean, J. Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781, 2013
57. https://www.datacamp.com/community/tutorials/stemming-lemmatization-python

58. https://medium.com/@athif.shaffy/one-hot-encoding-of-text-b69124bef0a7
59. J. Xue, E. Zhu, Q. Liu and J. Yin. Group Recommendation Based on Financial Social Network for Robo-Advisor. IEEE Access, Vol. 6, pp. 54527-54535, 2018
60. Ravi V, et al. Analytics. http://www.idrbt.ac.in/assets/publications/Staff%20Papers/Analytics.pdf, **3**(2), December 2017
61. Moody's Analytics
62. Kuo Wang, Ajay Kumar. Cross-spectral iris recognition using CNN and supervised discrete hashing. Pattern Recognition. Vol. 86, pp. 85-98, ISSN 0031-3203, https://doi.org/10.1016/j.patcog.2018.08.010, 2019
63. Nguyen, K., Fookes, C., Ross, A., & Sridharan, S. Iris recognition with off-the-shelf CNN features: A deep learning perspective. IEEE Access, 6, 18848-18855, 2017
64. Liu, N., Zhang, M., Li, H., Sun, Z., & Tan, T. Deep Iris: Learning pairwise filter bank for heterogeneous iris verification. Pattern Recognition Letters, 82, 154-161, 2016
65. Q. Zhang, H. Li, Z. Sun and T. Tan. Deep Feature Fusion for Iris and Periocular Biometrics on Mobile Devices. IEEE Transactions on Information Forensics and Security, **13**(11), pp. 2897-2912, doi: 10.1109/TIFS.2018.2833033, November 2018
66. Shuping Zhao, Bob Zhang, C.L. Philip Chen. Joint deep convolutional feature representation for hyperspectral palmprint recognition. Information Sciences, Vol. 489, pp. 167-181, ISSN 0020-0255, https://doi.org/10.1016/j.ins.2019.03.027, 2019
67. Sundararajan, K., & Woodard, D. L. Deep learning for biometrics: a survey. ACM Computing Surveys (CSUR), **51**(3), 65, 2018
68. Sun, X., Zhang, P., Liu, J. K., Yu, J., & Xie, W. Private machine learning classification based on fully homomorphic encryption. IEEE Transactions on Emerging Topics in Computing, 2018
69. Bost, R., Popa, R. A., Tu, S., & Goldwasser, S. Machine learning classification over encrypted data. In NDSS, Vol. 4324, pp. 4325, February 2015

70. Rajarshi Pal, Sastry VN, et al. Biometrics. www.idrbt.ac.in/assets/publications/Staff%20Papers/Biometrics.pdf. **3**(1), October 2017
71. K. Gunasekaran, J. Raja & R. Pitchai. Deep multimodal biometric recognition using contourlet derivative weighted rank fusion with human face, fingerprint and iris images. Automatika, DOI: 10.1080/00051144.2019.1565681, 2019
72. Mahfouz, A., Mahmoud, T. M., & Eldin, A. S. A survey on behavioral biometric authentication on smartphones. Journal of Information Security and Applications, 37, 28–37, 2017.

Recent Advances in Blockchain Technology

– **Dr. N. V. Narendra Kumar,**
Assistant Professor, IDRBT

Abstract

Blockchain technology is widely regarded as a disruptive technology capable of changing the very fabric of the way we connect and interact. Naturally, there has been a lot of effort dedicated to leverage the benefits of blockchain technology. Despite all the efforts, the technology is still in a nascent stage with a lot of refinements and fine-tuning still possible. In this article, we will explore some of the recent advances in blockchain technology.

1. Introduction to Blockchain Technology

1.1. Blockchain Technology

Blockchain is a shared and distributed ledger that keeps record of transactions. Blockchain by definition is append-only linked list data structure of blocks each containing a set of transactions. Each block is connected to its previous block by storing the hash of its previous block making a chain kind of structure. This type of data structure could be useful in a distributed environment where multiple entities generate data independently and still maintain the integrity of data.

In a distributed system, multiple entities transact with each other, and wish to maintain a verifiable and trust-worthy record of all the transactions without the need for a trusted central party. In general, a transaction workflow in a blockchain technology starts with the submission of a transaction to the network by an entity of the system. The submission of a transaction could be initiated by a human or by a smart contract, which is a computer encoding of the business logic. The transaction is validated against the current state of the system (denoted by the chain of blocks) and added to the new block. The new block is added to the chain of blocks after consensus is reached. Blockchain technology is realized by a combination of the following

components: (i) cryptography, (ii) peer-to-peer networking, (iii) distributed consensus, and (iv) fault-tolerant computing.

In distributed systems, one of the important aspects is to make the entities arrive at consensus on the list of transactions being added to the blockchain. This has reinvigorated research on distributed consensus models.

1.2. Characteristics of Blockchain Technology

Some of the basic features/characteristics of the blockchain technology are:

Immutability: Immutability can be defined as the property of an object of not being able to change its structure due to changes in its environment or external factors. It is impossible to achieve true immutability in digital systems. In the jargon of blockchain technology, immutability is used to mean that it is computationally infeasible to corrupt the data stored on the blockchain [1]. This is achieved by the use of hash functions, and by including the hash of a block **in its** successor block. By the nature of hash functions, blockchain makes it easy to verify and very difficult to modify the data stored in it, thus making it almost immutable.

Privacy: Preserving the privacy of data shared between the parties of a transaction is a critical requirement. In permissionless blockchains, all the data is shared to all the members of a network. However, the identity of the entities is pseudonymous. Thus, indirectly they provide limited privacy, but are prone to linking attacks. In contrast, in permissioned blockchains, because the data is shared only on a need-to-know basis, privacy is preserved better.

Consensus: The distributed consensus problem requires agreement among a number of entities for a single data value. Some of the entities may fail or be unreliable in other ways, so consensus protocols

must be fault-tolerant or resilient. In all the blockchain systems, achieving consensus on the state of the system is a crucial aspect and is realized through a variety of approaches. Permissionless blockchains use game theoretic strategies based on economic incentives, while permissioned blockchains use more traditional consensus strategies such as Paxos, PBFT, etc.

Security: Blockchain uses powerful cryptography to give individuals ownership of an address and the crypto assets associated with it, through a combination of public and private keys. This solves the issue of stolen identity as addresses are not directly associated with user's identity. And due to the data structure of the blockchain which is connected block, tampering of a record requires changing all the blocks in a chain. Thus, blockchain technology provides good security for user's data and also the transactions data.

1.3. Benefits of Blockchain Technology

Blockchain Technology, though an emerging technology, has already disrupted the current businesses in a way that there could be complete transformation of the solutions for any business. Some of the benefits of using blockchain technology are Transparency, Business Continuity, Disintermediation, Trust, and Smart Contracts, which are explained below:

Transparency: Transparency can be defined as a property where all the parties involved in a transaction know exactly what actions are being taken on what data. In blockchain, all valid transactions are added to the ledger, which is made available to all the parties based on the configuration, thus providing the necessary transparency.

Business Continuity: Businesses have a lot at stake based on the service provided to their clients. This makes the availability of services as one of the top requirements for any business. Despite failure of

some components, the system is expected to support continuity of business as usual. Permissionless blockchains provide business continuity because several copies of the data are distributed at multiple geographies.

Disintermediation: Disintermediation means replacement of intermediaries by technology components. Permissionless blockchains being truly decentralized enable a higher degree of disintermediation. However, this comes at the cost of loss of scalability and performance.

Trust: The problem blockchain is trying to solve is how to establish a trustworthy record among mistrusting entities. The protocols embedded in blockchains and the uses of cryptographic components are designed in such a way that trust is enforced and easily verified.

Smart Contracts: A smart contract is a computer code intended to digitally facilitate, verify, and enforce the negotiation or performance of a business logic. Smart contracts automate the performance of credible transactions without third parties. Though smart contracts can also be integrated with traditional systems, the guarantees of integrity of data in the blockchain together with the fact that all the parties see the same data, makes blockchain platforms particularly amenable to leveraging smart contracts.

2. How Business Can Benefit from Blockchain Technology

2.1. Business Requirements

Information and communication technology has been an enabler for business, and embracing the rapid advances in the field has become imperative to remain competitive. Particularly, in the case of

transactions involving multiple businesses, there are some important requirements that the underlying technology is expected to fulfil. In this section, we highlight the requirements that we consider are crucial for B2B systems.

Guarantees on message delivery: A trusted and reliable messaging framework is at the heart of any digitally-enabled business. The communications technology needs to guarantee that all the messages will be delivered to the intended recipients.

Guarantees on results of business processes: Though the main purpose of business networks is to enable exchange of information, it is equally important that the information exchanged is acted upon by both the parties in an agreed manner. When the records of the participants differ from one another, reconciliation becomes necessary which is resource-intensive and hence an expensive process.

Business continuity: Despite the advancements, systems are prone to errors either due to technical reasons or for reasons beyond our control. Businesses need the ability to run smoothly in spite of the system failures. Fault tolerance and recovery are important requirements for business continuity.

End-to-end security: Complex business processes require interactions among multiple participants, where information is processed and shared from one party to another in multiple hops. In such a scenario, it is important that all the participants are accessing and acting on the same information, and have the same understanding about the state (success / failure / current status) of the transaction. This requirement is referred to as end-to-end security.

Privacy: For successful conduct of business, it is important that information is available to all the participants who need to act on it. At the same time, it is equally important that the privacy is preserved

i.e., sensitive business information be not available to participants who do not need it.

Data integrity: Data integrity is the assurance that information is trustworthy and not corrupted. Maintaining the integrity of data is absolutely crucial for businesses since it is an invaluable tool for decision making that directly impacts the sustenance.

Latency: Latency is the time taken to complete a transaction. To efficiently tackle the growing volume of transactions, businesses need systems that provide a very low latency.

Asynchronous processing: Ideally, businesses should allow users the convenience to transact simultaneously through multiple channels, and have the capability to correctly handle all the transactions in a seamless manner. Such asynchronous processing also enables higher throughputs.

2.2. Characteristics of Current Central Systems

Currently, business networks have a hierarchical topology and are formed by establishing messaging platforms that enable information transfer. There is a single trusted party through which all transactions are routed. To understand this better, let us understand how a simple fund transfer works.

Steps for funds transfer:

1. Sender initiates the transaction to send money to receiver by sending the request to his/her bank
2. Sender bank on receiving the request transfers the request to the central bank after verifying the necessary conditions
3. The central bank upon receiving the request from sender's bank, executes the request and sends the response to the sender bank and also to the receiver's bank

4. Receivers bank on receiving the message from central bank updates the account of the receiver, to reflect the current updated balance.

Most traditional business networks follow a similar approach. Though these systems have evolved over time, there are certain concerns that need to be addressed:

1. Single point of Failure
2. Fault Tolerance
3. Reconciliation Overhead.

Single point of Failure

A single point of failure (SPOF) is a potential risk posed by a flaw in the design, implementation or configuration of a circuit or system in which one fault or malfunction causes an entire system to stop operating. In a data centre environment, a single point of failure can compromise the availability of the system or the entire data centre depending on the location and interdependencies involved in the failure. Adding redundancy to such systems could solve the problem to an extent. However, it introduces the challenge of synchronizing the data between replicas.

Fault Tolerance

Fault tolerance is the property that enables a system to continue operating properly in the event of the failure of some (or one or more faults within) of its components. If the operating quality for a system decreases, the decrease will be proportional to the severity of the failure. Fault tolerance is particularly sought after in high-availability or life-critical systems.

Reconciliation Overhead

Reconciliation is an accounting process that uses two sets of records to ensure figures are correct and in agreement. It confirms whether the money leaving an account matches the amount that has been spent, and making sure the two are balanced at the end of the recording period. The purpose of reconciliation is to provide consistency and accuracy in financial accounts.

In summary, though the currently used central systems address the business requirements to a large extent, there are improvements possible. In particular, our study finds that central systems can be enhanced using the available technologies to alleviate their shortcomings in – (i) guarantees on business processes i.e. reconciliation efforts, and (ii) business continuity by adapting a replication strategy which increases redundancy with appropriate mechanisms for ensuring consistency of data. However, both these efforts introduce processing overheads, which impact the throughput. Further, we find that end-to-end security is difficult to achieve in current systems through simple enhancements.

2.3. How Blockchain technology can complement current technology

Blockchain technology can complement the existing business networks to overcome some of the limitations. Key features of the blockchain are: distributed ledger, near real-time updates, chronological and time-stamped data, cryptographically sealed, and programmable and enforceable contracts. These features can enable the business network to exploit blockchain technology for:

Guaranteeing Results of Business Processes: Distributed shared ledger along with programmable and enforceable contracts in blockchain provide this feature

Improving Data Integrity and Finality: Cryptographically sealed ledger with chronological and time-stamped transactions in blockchain provide this feature

Providing a Shared View: Near real-time updates coupled with distributed ledger provides shared view

Validating Business Rules: Programmable and enforceable contracts provide the mechanism to enforce and validate the shared business rules.

3. Innovations of FinTech in Blockchain Technology

As mentioned earlier, blockchain technology is still an emerging technology and currently there is a lot of research happening all over the world. Though blockchain technology has no production-ready systems, IT giants, banks and financial organizations are investing in blockchain research to unearth the potential use cases which could benefit from this technology. Many consortiums are also being established to design and develop solutions for many business cases. Research is being done on different aspects of blockchain technology. It is notable that many FinTech companies are contributing in advancing the technology by constantly innovating.

3.1. Blockchain Data Structure

The linked list data structure forms the core of blockchain technology. Though a simple data structure is sufficient to provide necessary functions, there are some innovations at data structure level described as below:

Hedera Hashgraph: Hedera Hashgraph is a DLT project and an innovative variation of a standard blockchain on the basis of the technology called directed acyclic graph (DAG) [4]. Basically, operations are recorded not via the chain (like in a blockchain), but via

a directed acyclic graph. This new DLT has been adopted in the Swirlds platform where the consensus protocols are different from the blockchain based consensus. Hashgraph uses virtual voting consensus mechanism, which as of now is not compatible with any other existing blockchain platforms like Hyperledger Fabric, and Multichain.

Tangle: Tangle is a new concept which has no mining, no transaction fees and is fast [12]. In Tangle, transactions are linked to each other like a big web tangle. There is no concept of blocks. Each transaction will be verified by a little proof-of-work by the originating device. It will link the transaction further to two more random transactions. In Tangle, users do not need to wait for blocks to be mined. The transactions are verified in parallel. This helps in fast and more number of transactions at a time. This data structure has been adopted by the IOTA blockchain.

3.2. Consensus/ Protocol

Consensus is the core part of any distributed technology as it provides the provenance of all the transactions executed in the platform. Many blockchain platforms use a simple consensus called 'Proof-of-Work' where the miner has to prove that the new block being added is the correct block containing correct set of transactions. A good amount of research is happening for designing and developing new consensus protocols, which are listed as Proof of Authority, Proof of Concept, Proof of Principle, Proof-of-Stake, delegated Proof of Stake, Proof of Mechanism, Proof-of-Space-Time, Proof of Elapsed Time (PoET) introduced in Hyperledger Sawtooth platform, Redundant Byzantine Fault Tolerance (RBFT) in Hyperledger Indy, Sumeragi in Hyperledger Iroha [13], Loop Fault Tolerance [14], Kafka in Hyperledger Fabric, Algorand consensus [15]. Apart from these consensus protocols, which work well for blockchain type of DLTs, virtual voting consensus protocol is being used for Tangle DLT.

3.3. Security

Security being one of the key characteristics of the blockchain technology, there has been an enormous effort to handle the latest cyber-attacks in the DLT systems. Current state-of-the-art of cyber security handles a good amount of attacks, but the area still requires advancements for future. Some of the advancements in security are based on Secure Multiparty Computation, Zero Knowledge Proofs and Secret Sharing Mechanisms.

Secure Multiparty Computation (SMC): SMC is defined as a way for multiple parties to run a computation on a series of inputs for which each party only knows one of the inputs [16]. For example, suppose that Alice and his group of friends want to know the average salary of the group without involving any strangers and also without revealing any individual's salary. To start with, the first person picks a random element, which is known only by him, to modify his salary – let us say he adds the random number 550,500 to his salary, and then shares the result of that addition with the person next to him. The next person then adds his salary to the first result, and passes the result of that computation on to the third person, and so on, until all the salaries have been summed and returned to the first person. The intermediate results are passed securely from the previous person to the next person, and only the final sum is returned to the first person. The first person then subtracts the random element 550,500 from the total, and then may calculate the average salary of the group without knowing the salary of any one individual. In practice, SMC is generally far more complex than this for a variety of reasons (such as to prevent collusion amongst the parties involved).

Zero-Knowledge Proof: A zero-knowledge proof provides a witness of some valid statements without revealing information beyond the validity itself. Proofs, which require no trusted setup and offer efficient proof construction and verification, can be combined with Ethers to achieve zero-knowledge proofs in Ethereum. In Blockchain,

zero-knowledge is implemented as zk-SNARKS[6][7] which are used by the privacy-preserving Blockchain platforms. In such systems despite information being published on a public blockchain, the sender, recipient, and message can remain private. It is also incorporated in Quorum Blockchain, and is currently available in Ethereum and Hyperledger Fabric.

Indistinguishability Obfuscation

Indistinguishability obfuscation (IO) is a cryptographic primitive that provides a formal notion of program obfuscation. Informally, obfuscation hides the implementation of a program while still allowing users to run it. A candidate construction of IO with provable security under concrete hardness assumptions relating to multilinear maps was published in 2013.

Typical blockchain design is centred around a widely distributed, global database which stores all transactions that have ever taken place in the system. Thus, there is no avenue for redress if a user wishes to retrospectively hide a transaction. Further, nothing in the ledger is encrypted, and digital signatures are mandatory, ensuring cryptographic attribution of activities to users. On the other hand, account identifiers in blockchain platforms take the form of cryptographic public keys, which are pseudonymous. Anyone can use their application interfaces to trivially generate a new public key and use it as a pseudonym to send or receive payments without registering or providing personal information. However, pseudonymity alone provides little privacy, and there are many ways in which identities could be linked to these pseudonyms. To counter this, many platforms employ a variety of obfuscation techniques to increase their financial privacy.

Among the techniques used in blockchain, the most prevalent can be characterized as “ambiguating obfuscation” [17]: effectively reducing the information an adversary is able to extract from a particular

transaction. Examples include using a new pseudonym for every new transaction and randomizing the structure of transactions to make the spend to the "true" recipient indistinguishable from "change" going back to the sender. A second type of obfuscation, namely "cooperative obfuscation", has risen in popularity over the last years. For example, users can send their money to a service that will "mix" their funds with those of other users, thereby obfuscating the flow of payments. A similar technique called CoinJoin [18] works in a peer-to-peer fashion and does not require a trusted intermediary. Due to the need for these users to find and transact with each other, markets for anonymity preserving solutions have emerged that bring together providers and receivers of anonymity.

Trusted Execution Environments (TEE)

The TEE is an isolated environment that runs in parallel with the operating system, providing security for the rich environment. It is intended to be more secure than the user-facing OS and offers a higher level of performance and functionality than a Secure Element, using a hybrid approach that utilizes both hardware and software to protect data. It offers a level of security sufficient for many applications. Only trusted applications running in a TEE have access to the full power of a device's main processor, peripherals and memory, while hardware isolation protects these from user installed apps running in a main operating system. Software and cryptographic isolation inside the TEE, protect the trusted applications contained within from each other. Some of the early research, related to TEE, are Intel's SGX and TrustZone from ARM.

> **Intel's SGX**: Intel's Software Guard Extensions (SGX) [2] is a set of extensions to the Intel architecture that aims to provide integrity and confidentiality guarantees to security sensitive computation performed on a computer where all the privileged software (kernel, hypervisor, etc.) is potentially malicious. SGX relies on software attestation, which proves to

a user that she is communicating with a specific piece of software running in a secure container hosted by the trusted hardware. The proof is a cryptographic signature that certifies the hash of the secure container's contents. It follows that the remote computer's owner can load any software in a secure container, but the remote computation service user will refuse to load her data into a secure container whose contents' hash does not match the expected value.

TrustZone from ARM: ARM processors with TrustZone implement architectural Security Extensions in which each of the physical processor cores provides two virtual cores – Non Secure World and Secure World, and a mechanism to context switch between the two, known as the monitor mode [19]. The distinction between both worlds is completely orthogonal to the normal ring protection between user-level and kernel-level code and hidden from the operating system running in the normal world. When secure mode is active, the software running on the CPU has a different view on the whole system than software running in non-secure mode. This way, system functions, in particular security functions and cryptographic credentials can be hidden from the normal world. It goes without saying that this concept is vastly more flexible.

Blockchain Open Ledger Operating System (BOLOS)

BOLOS is a new operating system [10] which represents a major change compared to what the smartcard industry offers today. It puts developers in the driving seat, providing an unobtrusive framework to build source code portable native applications around a secure core, protecting the core against applications attacks, and isolating applications from each other without getting in the way. BOLOS is a way of turning Hardware Wallets into Personal Security Devices where users can review and install third party applications that will add new privacy features on top of their own shared set of cryptographic material, without exposing that material.

3.4. Smart Contracts

Smart contracts are computer code/software that automatically executes transactions (e.g., exchange of money, property, shares or anything of value) and/or enforces agreements based on the fulfilment of the terms of an agreement. Ethereum was the first blockchain platform to introduce the smart contracts written in a language called Solidity. DAML [3] is a Digital Asset Management Language to write generic smart contracts.

4. Application Domains of Blockchain Technology

In the previous section, we have discussed the innovations in the core Blockchain technologies. This section presents the application domains in which the benefits of Blockchain technology are being leveraged to disrupt the industries.

4.1. Finance

Blockchain as a Service

Blockchain as a service is an offering, which enables the customers to leverage the cloud-based solutions for building, hosting, and using their smart contracts, blockchain applications, and several other blockchain functions. In the case of BAAS [11], the cloud service provider manages all activities and tasks to keep the infrastructure operational and agile.

Enterprises like IBM, Oracle, and Microsoft, etc. have already started offering Blockchain-as-a–Service, thereby enabling enterprises to be involved with the technology without risking a high investment cost that would go into in-house blockchain development.

Asset Tokenization

The process of asset tokenization entails using digital tokens to prove the ownership of real assets. It works with the primary purpose of expediting the buying and selling process of security, thus becoming one of the biggest uses of blockchain technology. The various benefits that tokenization of assets offer – greater liquidity, faster transactions, more accessibility, and faster transactions, have all together attracted a number of players to enter the blockchain application development market surrounding Asset Tokenization.

IoT Inclusion

The decentralized market has found several use cases for merging IoT with blockchain based applications. The fact that passing data with utmost security lies at the core of blockchain technology fits almost naturally with the mechanism of IoT that drives on connecting devices and transferring information between one platform to another.

Blockchains Interoperability

There has been a need to have a solution that helps organizations to interoperate between the available blockchain platforms like Hyperledger Fabric, Corda, Multichain. Some of the solutions available for this purpose are Cross Chain Technology, Federated Blockchain and Hybrid Blockchains.

4.2. Artificial Intelligence

Artificial intelligence and blockchain are the two most talked-about tech trends in the past two years by a wide margin. AI promises to automate many tasks and is often better at modelling complex situations than humans. Blockchain, on the other hand, offers greater data security and privacy while also reducing the overhead and centralized power of major institutions.

Energy Industry

Applications combining Blockchain and AI [8] are being developed to help some of the world's largest energy companies conceptualize and more effectively use data. Analytics on energy data, anonymously collects data from multiple locations/companies and uses smart contracts to give businesses better overall insights into the energy industry.

AI-driven Trading

AI-driven trading platforms are being developed to give traders AI-based scoring systems, high-quality data feeds and advanced forecasting, which can provide personalized news and trading features to maximize their efforts.

AI-Entertainment

Blockchain-powered entertainment platform [20] that uses AI to enhance audience analytics and viewership algorithms is being designed, with the objective of making it available for public consumption. The combination of AI and blockchain powers the "Proof-of-Engagement" (POE) tools that automatically trigger payments to content providers based on views and engagement times.

4.3. Healthcare and Pharmaceuticals

Healthcare and Pharmaceuticals industries have been working on leveraging the blockchain technology to address various use cases [21], some of which are detailed below:

Verifying the authenticity of returned drugs

Verifying the authenticity of the drugs, which are returned from the retailers and the wholesalers to the industry is very much needed as these return drugs itself, constitute 2-3 percent of sales costing

around $7-8 billion. By leveraging the blockchain characteristics, it is easy to track down the unauthenticated drugs.

Prevention of counterfeit drugs and medical devices

Blockchain's ability to establish provenance of data, makes it especially suited for this supply chain use case. Drug companies that manufacture, ship and deliver products have a difficult time keeping track of their products, thereby allowing counterfeiters to introduce fake drugs into the system. The problem of counterfeiting is not just limited to drug manufactures, but extends to medical instruments manufactures.

Compliance in pharma supply chain

As drugs move through the supply chain, logistics companies need to adhere to drug handling, transport and storage guidelines. Several operating constraints may need to be tracked. Blockchain technology provides the pharmaceutical supply chain a better way to add compliance and governance within the supply chain.

Transparency and traceability of consent in clinical trials

A patient consent should be informed and aware of each step in the Clinical Trail Process. Blockchains provide a mechanism for unfalsifiable time-stamping of consent forms, storing and tracking the patient consent in a secure, and publicly verifiable way enabling the sharing of this information in real-time. Blockchain protocols in clinical trials can provide transparency and traceability of consent.

Improving the quality and reliability of clinical trials data

Patients and physicians have begun to question the current standards for Clinical Trial funding, leading to growing concerns due to conflict of interest and the high stakes involved. Blockchain with its decentralized,

immutable ledger and a mechanism to ensure transparency provides a solution.

4.4. Insurance

While blockchain might not be the end-all-be-all to problems faced by insurers, it does provide foundational technology that promotes trust, transparency and stability. Blockchain is in the early stages of adoption, but there are already a handful of ways that insurers are leveraging the technology to mitigate the below-mentioned challenges [9].

Enhance efficiencies

Blockchain provides a solution to drive efficiency and security that would allow the personal data to be controlled by an individual while verification is registered on the blockchain. Some companies are trying to develop blockchain solutions for know-your-customer (KYC) data. The goal would be to have the KYC data verified and then it could be securely forwarded to other companies to use without the need to repeat the data entry or verification process.

Improved claims processing through smart contracts

The insured and the insurer each currently have issues that blockchain and smart contracts could solve. Insured individuals typically find insurance contracts long and confusing while the insurance companies are battling an extraordinary amount of fraud. Through blockchain and smart contracts, both parties would benefit from managing claims in a responsive and transparent way. It would start by recording and verifying contracts on the blockchain. When a claim is submitted, the blockchain could ensure that only valid claims are paid. The network would know if there were multiple claims submitted for the same accident. When certain criteria are met, a blockchain could trigger payment of the claim without any human intervention, therefore improving the speed of resolution for claims.

Fraud detection and prevention

Validation is at the core of blockchain technology's decentralized repository and its historical record, which can independently verify customers, policies and transactions for authenticity. In order to work to its full potential, this would require extensive cooperation between insurers, manufacturers, customers and other parties who would use the blockchain to share info to prove policies, purchases of products, verify police reports and more.

4.5. Cyber Security

Innovative uses for blockchain technology are already becoming a part of other fields and can be especially useful to boost cyber security.

Minimizing data destruction

The blockchain could play an important role in guaranteeing data availability, because unlike a typical scenario, every piece of information will be distributed throughout the entire system. This means that the only way that the data can be accidentally or maliciously destroyed is for the whole blockchain to be wiped out, and then wipe out every single node separately. If even one node remains within the system, the data could be fully restored.

DNS distribution and DDoS prevention

Distributed Denial of Service (DDoS) attacks have recently become more frequent and potent for big companies [23]. With blockchain, Domain Name Server (DNS) would be completely decentralized, so that each DNS would point to multiple nodes. In this scenario, an attacker would have to direct a DDoS attack to all nodes in the blockchain, making it almost impossible for a website to be taken offline. Some companies have emerged with an implementation of decentralized DNS in blockchain, so as to prevent DDoS attacks from occurring.

Transparent and Immutable

Blockchain can be seen as a self-auditing ecosystem of some digital value where the network reconciles every transaction that happens in some intervals. Being able to access all the blocks over the network provides more transparency. And due to the immutable characteristic of blockchain, to change any single transaction would require computational power to change the entire Blockchain of the system.

4.6. Governance

IT giants have been working on a blockchain transformation for the governments. They are providing solutions such as cross-border transactions, education and so on.

Cross-border Transactions

Any government would strive for the betterment of its economy. Since trading is a major source of economy, it would be helpful to have digital platforms that enable hassle free cross-border trading. Having blockchain enabled cross-border transaction would make the trading free and more transparent. Many governments have taken initial steps to take the make progress in this direction.

Education

Some of the use cases where blockchain can be used in education domain are:

Diplomas and certificates: The academic record of the students, which contains notes, diplomas and titles obtained, is protected. This information is available in the chain of blocks even in case the institution loses these files. It also allows security to be reinforced so that the diplomas are not modified.

Security in the archives: In virtual education, sometimes, the theft or plagiarism of investigations or projects is presented by those who try to present a new one. Also, the documents or files prepared by

students and institutions will be safe with this technology, to avoid theft in both cases and forgery of digital signatures.

Reliable transactions: With the blockchain, it is very practical to manage economic transactions with online institutions or study centers. Also, one can verify the credibility of online institutions, and thus can help in avoiding "falling into the trap" of fraudulent networks.

Accreditation of credentials: It focuses on interpersonal skills and it is possible to obtain them through "peer-to-peer" processes, better known as P2P. It is based on the fact that the students with whom the projects are carried out can give credentials to their peers, certifying some skill developed during the group learning. What you get is known as an "open badge," which can support the knowledge and skills acquired during the course or virtual learning.

4.7. Waste Management

There are a number of waste management sectors for which blockchain based projects could bring some real benefit. A notable addition is in the household and industrial waste sector.

Waste Sharing Initiatives

The current model of waste management means – we centralize trash in landfills. This creates a concentration of waste, which produces methane as a by-product. Methane is more harmful to the environment than carbon dioxide, up to 25 times in fact. If we were to distribute our waste among the population, we would be able to achieve a couple of key benefits: the waste would decompose quicker, and we could create fertilizer for local gardening projects. Blockchain for the environment could be not necessarily just tech-based solution but can provide a platform to develop recycling efforts [24].

5. Conclusion

Blockchain technology has a great potential to revolutionize the way we conduct business. However, several factors are creating roadblocks for successful implementation of this technology in practice. There has been good progress in solving many important challenges for adopting blockchains for business. FinTechs have been actively contributing to the development of this technology both through the innovations in the core aspects itself and through its application to innovative use cases.

References

1. K. Sultan, U. Ruhi, R. - Lakhani (Eds.). Conceptualizing blockchain: characteristics & applications, 11th IADIS International Conference Information Systems. pp. 49–57, 2018
2. V. Costan and S. Devadas. Intel SGX explained. Cryptology ePrint Archive, Report 2016/086. 2016, http://eprint.iacr.org/
3. Digital Asset. The Digital Asset Platform – Non-technical White Paper, 2016. Available at https://digitalasset.com/press/digital-asset-releases-non-technical-white-paper.html.
4. Dr. Leemon Baird, Mance Harmon, and Paul Madsen. Hedera: A Public Hashgraph Network & Governing Council. White Paper v.1.5, 2019
5. Dean Demellweek. Blockchain-based Zero Knowledge Proof solution (Link). July 2017
6. Koens, T., Ramaekers, C., & van Wijk, C. Efficient Zero-Knowledge Range Proofs in Ethereum. November 16, 2017. Retrieved from https://www.ingwb.com/media/2122048/zeroknowledge-range-proof-whitepaper.pdf
7. Christian Reitwiessner. zkSNARKs in a nutshell (Link). December 5, 2016
8. Raghav Bharadwaj. AI in Blockchain – Current Applications and Trends (Link). May 19, 2019
9. Inmediate.io. The Potential Of Blockchain Technology In The Insurance Industry (Link). November 8 2018

10. Introducing BOLOS: Blockchain Open Ledger Operating System (Link). March 2016
11. Z Ming, S Yang, Qi Li, D Wang, M Xu, Ke Xu, Laizhong Cui. Blockcloud: A Blockchain-based Service-centric Network Stack, 2018
12. Serguei Popov. The Tangle (http://iotatoken.com/IOTA_Whitepaper.pdf), 2016
13. Christian Cachin. Architecture of the Hyperledger Blockchain Fabric. IBM Research – Zurich CH-8803 Ruschlikon, Switzerland. July 2016
14. ICON Foundation. Hyperconnect the World. January 2018
15. Y. Gilad, R. Hemo, S. Micali, G. Vlachos, and N. Zeldovich. Algorand: Scaling byzantine agreements for cryptocurrencies. in Proc. 26th ACM Symp. Operating Syst. Principles. pp. 51–68, 2017
16. Oliver Birch. Secure Multiparty Computation and Shamir's Secret Sharing on Wanchain (Link). July 20, 2018
17. Arvind Narayanan & Malte Möser. Obfuscation in Bitcoin: Techniques and Politics, International Workshop on Obfuscation: Science, Technology, and Theory, New York University. June 2017
18. Fahad Saleem. What is CoinJoin: Detailed Explanation for Beginners (Link). 2019
19. Scott Thornton. Arm Trust Zone explained (Link). December 28, 2017
20. https://www.alphanetworks.com/en
21. Prashant Ram. Real world Use Cases of Blockchain in Pharma and Healthcare (Link). August 2018
22. DQINDIA Online. The Inevitable Blockchain Tomorrow of Cybersecurity, Governance and Law (Link). June 2019
23. Claudio Buttice. Will Blockchain Technology Make DDoS Attacks Obsolete? (Link). January 15, 2018
24. Recereum. Solving the Global Waste Problem using Blockchain Technology (Link). 2018.

* The contribution made by Shri. Ravi Kanth Kotha, Research Fellow, IDRBT is gratefully acknowledged.

Emerging Security Solutions

- **Dr. V. Radha,**
Associate Professor, IDRBT

Abstract

The two major factors that are driving new security offerings in the BFSI are: 1. Regulatory requirements and 2. Increase in targeted attacks. Since targeted attacks need targeted defence, new solutions must be adapted. This paper describes the latest trends in the prevention of cyber-attacks.

1. Introduction

Cyber security is a major concern for any business and especially the banking and financial sector. As the financial sector deals with the business of money apart from critical data, it is a jackpot for the hackers. The age-old security solutions to protect the data may not be sufficient, as hackers continuously upgrade their tactics. With the changing threat landscape, even businesses need to update their security strategies. In this paper, we discuss a few solutions for cyber security, which benefits the financial sector.

1.1 Current Cyber Security Threats in Banking and Financial Sector

The major security threats, especially to the banking and financial sector comes from:

- **Unencrypted data:** Majority of the data breaches occur due to no proper encryption and the stolen data is always available immediately without much effort
- **Automation technologies without proper security:** Attackers can turn CCTV cameras, connected devices, and other electronic gadgets into bots, which can cause a lot of damage
- **Third party services:** The internet connects everything, so unprotected third party services can open many roots for cyber-attacks

- **Unsecured mobile banking:** Most of the mobiles now have the feature of mobile banking, which leads attackers to take advantage of it. This necessitates one to offer mobile apps with lots of in-built security
- **A constantly changing threat landscape:** Attackers and their motives keep changing day-to-day. Therefore, the approach to protect assets also must change accordingly
- **The Big Breach:** The fear of big breach keeps CISOs awake at night. CISOs are under huge pressure to keep customer data safe from hackers and fraudsters
- **Different forms of hacking:** These days attackers do not rob the important data, but keep it hostage by releasing different ransomware each day
- **Customers' account takeovers:** Customer accounts are hacked almost every day in some place or the other. Hence, spreading customer awareness about such thefts and a word of caution is important.

2 DLP (Data Loss Protection)

DLP is also known as Data Loss Prevention/Protection; Data Leak Prevention/Protection; Information Loss Prevention/Protection; Information Leak Prevention/Protection; Extrusion Prevention; Content Monitoring and Filtering; Content Monitoring and Protection.

Definition of DLP: Data loss prevention (DLP) is a method used for making sure that the employees within the company do not send / leak sensitive/critical or important information to the outside world. DLP helps a network administrator to control what data the employees can transfer.

Data loss is one of the biggest threats to the financial sector and in most cases, it is due to malicious insiders.

Customer data flows through many points within a bank's IT infrastructure. The complex and dynamic nature of today's banks' IT system makes things even more difficult. At times, sensitive data is vulnerable due to employees' accessing their emails and sending info to the outside world. A complete DLP solution will monitor all communications, cloud, external devices for access or manipulation of sensitive data by users.

To implement DLP controls effectively, an organization must answer these three fundamental questions: 1. What sensitive data does it hold?; 2. Where does its sensitive data reside, both within and outside the organisation? and 3. Where does its data move to?

DLP is a Process, not a Product

For a security program to be mature, DLP must always be an integral part of it. However, it is very important to note that it is not a product, but a process. In order for DLP to be successful, organizations need to plan, control and chart out polices accordingly.

We need to understand the difference between content awareness and contextual analysis. We can assume that context is an envelope and content is analysing what is inside it. Content discovery is an important phase of any DLP solution.

Content Discovery Techniques

There are three basic techniques for content discovery:

a) **Remote Scanning:** Solution server performs remote scan on the systems
b) **Agent-Based Scanning:** The agents installed, perform a scan
c) **Memory-Resident Agent Scanning:** Instead of deploying a full time agent, we can deploy a memory resident agent, which performs scan and then exits without leaving any footprint.

We discuss different DLP solutions, each focused on a specific purpose, but with the same goal, which is to prevent data loss.

2.1 DLP on Network

The network DLP solutions can stop data leakage that may happen through an organization's network communications like web applications, email and FTP. Typically, the network DLP solution monitors the traffic and prevents data loss regardless of the port or protocol; it can also encrypt certain content like card data before transmitting; it can notify both the users and admin when the network traffic violates the policies present.

These solutions cover the most common communication mediums/protocols like HTTPS, HTTP, IMAP, SNMP and POP3. Here, the focus is in protecting the entire organizational network. However, these solutions neither offer protection from external devices nor do they restrict copy and paste options of files on the employees' workstations. They install a proxy or sniffer at application level.

The key features one should look from while choosing a network DLP includes: automatically blocking or warning users if an activity looks risky, encrypting all the data that is leaving the organization, logging everything for further incident response and forensic analysis.

2.2 DLP on Storage

The storage DLP allows one to classify the users and confidential files (tag/label); then gives access permission as per label to the groups of employees. Hence, it is possible to find out the sensitive points and prevent data leakage. This is a good solution for data stored on premise as well as in the cloud.

The key features one should look for while choosing a Storage DLP includes: integrating with existing cloud storages to prevent sensitive

data leak, scan the data uploaded and continuously audit existing data.

2.3 Endpoint DLP

In the initial days, disks and in the recent times, pen drives are the external devices or tools, which is the reason for leaks from PC based systems, laptops, tablets, etc. They may be practically very fast in transferring files and data, but they put the organizations' security at risk and become one of the prime reasons for intentional or accidental data leakage.

In this case, the solution has central management that controls rules applied on the client side devices. (The agents are installed in the client side machine and then managed centrally). In these solutions, copy and paste options, screenshot grabbing options are disabled.

The key features one should look for while choosing an endpoint DLP includes: automatic logging and warning when an external device is connected.

Best Practices for DLP

1) Always create a document classification matrix before implementing the solution. This practice will classify the documents according to the risks, with public data given high priority
2) Start user education to obtain employees confidence and to be more careful. It is important to ensure protection of both the newly created and existing data
3) Identify the store where the data at present resides and its classification. Learn from peers and top management to understand how data flows within and without the network. Examine controls and data stores currently in place
4) Identify top data loss scenarios. These controls reduce the risk of data loss and lessen the damage of such data loss scenarios

5) People working with highly sensitive data must be provided with endpoint DLP agent, so that no data is leaked, copied and pasted or screenshot-grabbed
6) Create simple policy rules; always start with monitor mode. Monitor mode is preferred as it provides insight into how to enforce policies
7) Use network based DLP for monitoring access to sensitive files and for the knowledge of current data usage within the company and the data flows within the network and internet
8) A risk-based approach must be followed for blocking threats. The blocking can be classified as allow, log only, block and log. To avoid false positives, the blocking capacity must be increased by strengthening the policies
9) The organization must allow users to correct their mistakes in real-time i.e. modify their documents in order to meet the policies
10) With reviewing logs and modifying policies, the organization must also warn the key offenders by sending warning mails or so
11) Gap Analysis must be conducted to see the current risk level for data loss and the acceptable risk level.

Different Softwares for DLP

Table 1:

Solution	Capabilities	Autonomous features	Cloud compatible	Remarks
Symantec	Data management and tracking	No	Yes	Cloud Coverage; Very much enterprise-oriented
Trustwave	“”	Yes	Yes	Automatically blocks threats; too many settings
McAfee	“” + Forensic	“”	“”	Intelligently

				prioritizes data
Digital Guardian	Data management + tracking + encryption	“”	“”	Can be on premise, cloud or hybrid
Check Point DLP	Data loss education and remediation	“”	“”	Overly simple to use

3 Deception Technologies

Usually in cyber security, the attackers have an advantage of succeeding with just one attack and with an in-depth knowledge of few vulnerabilities, but the defenders have to be more aware of all vulnerabilities, platforms, networks, etc. to defend against the attackers.

These technologies gained importance to make the life of attacker difficult. Decoys are used to misdirect the attacker and delay or prevent her/him from going deeper into the network and reaching his intended target.

Definition:

- Deception technologies can detect, defend and analyse against zero day and advanced attacks in real-time
- The deception technologies gives us insight into malicious activity within internal networks, which otherwise were ignored by traditional cyber defence
- This technology gives us a rare advantage over the attackers by doing something different. They provide an early and accurate detection by planting attractive decoys and content to attract attackers. This can all be done within the company’s network and it serves as a warning system for attacks that have been bypassed through the security perimeter.

Types of Activities Deception System Defends:

Deception systems detect a wide variety of threats and are not dependent on known signatures, databases look up, or pattern matching:

Credential Theft	When somebody tries to attack and lift username and passwords from OLAP directories or other places they are stored.
Lateral movement	When someone in one part of network tries to access other parts of the network that are off limits to them.
Attacks on directory systems	These can be directories of users or file directories.
Man-in-the-middle	These are attacks where a hacker intercepts and possible changes communication between two parties who do not know their exchanges have been infiltrated.
Sensitive data	This is sensitive, high-value data.
Geo-fencing	The attacker steals planted deception files that when opened provide geo-location data and intelligence on targeted files.
Detecting attacks on data distributed outside your organization	It is often possible, using decoy docs, to put essentially GPS trackers on your data to see when it's been accessed even after it has left your system.

Advantages of using Deception Technology

- Deception reduces the dwell time and mean time to detect and remediate
- By engaging the attacker, it provides insight into forensics of adversary intelligence including indicators of compromise (IOCs) and tactics, techniques, and procedures (TTPs)
- If we suspect that an attack may happen then we can automatically add few more decoys around our critical assets or simply reset the attack surface, so that we can up the complexity for an adversary
- Deception technology is capable of finding attacks like advanced persistent threats (APT's) and Zero Day Events that the endpoint protection systems deployed failed to stop

- Deception techniques can catch and help prevent major attacks with high accuracy and in time, reducing the risk of economic loss due to data theft and improving overall business operations
- Deception techniques are capable of revealing attacker activity and intentions as they can provide comprehensive visibility into internal networks
- Integrations with current defence in depth solutions and security operations solutions are available and an ongoing attack can easily be terminated with such integrations.

Financial organizations can use deception-based threat detection for the purposes listed below:

Derailing Attacks: Financial institutions use deception to complicate, slow down, and deviate the attacks with the help of decoys that can detect early in-network attacks and attack movement within the network. Attractive deception will lure the attacker by intentionally leaving credentials and data, and misdirect attacks into an engagement server, which raises a high-fidelity alert.

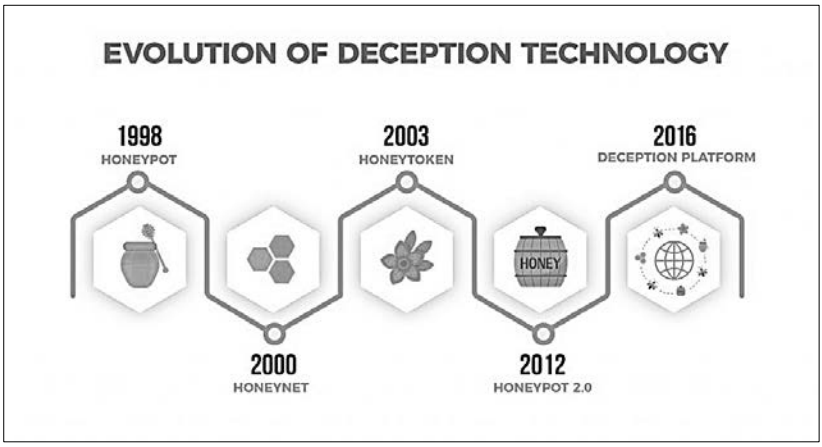
Visibility to Lateral Movement: The deception makes attackers reveal themselves as they keep trying to attack or as they start laterally moving across network services, virtual machines, IP service, and subnets looking for high-value data assets.

High Fidelity Alerts: Used to raise alerts based on the engagement of an attacker. The alert includes the threat intelligence and forensic reports on affected systems, attacker's activity, and signatures required to act decisively and quickly. Integrations are available for easy threat intelligence sharing and automation.

Proactive Defence: This provides an ability to increase the attack surface by adding decoys on the go, imitating production environment. This makes the job of attacker more difficult, increases the cost and forces him to make a mistake that reveals his presence.

3.1 Honeypots and Honeynets

Deception Technology has emerged based on honeypot, honeynet and honeytoken technologies.



Source: https://www.wwt.com/all-blog/deception-technology/

In computer security, a honeypot is a trap, which is set to detect attempts at any unauthorized use of information systems.

Generally, a honeypot consists of data that appears to be a legal/correct part of the site, but is actually isolated and watches upon attackers when they assume that data to be actual and attack the honeypot. A honeypot diverts the attention of the attacker from the main network and can identify new vulnerabilities and risks.

A group of honeypots together forms honeynet. It is an architecture one can populate with live systems, and is not a product or software. Any traffic entering or leaving is suspect. Each packet entering the honeynet is monitored, captured and analysed.

Honeypots in future are going to become a critical weapon in security services as they have the capability to catch new hacker methodologies, toolkits and scripts. Banking sector would further benefit because data is monitored in real-time and requires minimal hardware.

3.1.1 Honeypots based on interaction level

- **Low-interaction honeypots:** In these honeypots, we simulate the services that are frequently targeted by attackers. Since they consume relatively few resources – one physical machine can host many virtual machines – their response time is shorter (quicker) with minimal coding, thus, reducing the complexity of the virtual system's security like web applications.
- **High-interaction honeypots:** These honeypots copy or imitate the exact activities of the production system, which host a wide range of services, and therefore the attacker may waste a lot of time thinking that they are the original resources. Multiple virtual machines can be made honeypots and can be hosted on a single physical machine like honeynets.

There are many other types of honeypots like pure honeypots, medium interaction honeypot, malware honeypot, spam honeypot, database honeypot, etc., which can be used based upon the environment of the organization.

Things to consider before building own honeypot:

- Logging must be done for all packets going and coming from a honeypot system
- To analyse the attacks, a protocol analyser must be used
- Whenever there is heavy traffic from honeypots, we can utilize the firewall notification feature to send alerts to us, so that proper logging and monitoring can be done
- To watch external cyber-attacks to our network, the best way is to place honeypots outside the network and towards the internet.

3.2 Best Practices of Deception

- **Deception must be authentic:** For the deception to work, the decoy we deploy must seem very real, so that, we can lure the attacker in and learn about his intensions and methodologies
- **Deception must be comprehensive:** An efficient Deception Technology must cover an ever-changing attack surface by adding decoys dynamically
- **Deception must be scalable:** Deception is more efficient than other means of detecting cyber security threats because it does not analyse traffic or behaviour or does database lookup, etc., which are memory- and process-intensive, and raise high false positive alerts. With good understanding of the IT and business environment, and appropriate placement of decoys, Deception Technology is easy to scale and gives better results.

Deception technology vendors and tools

One can implement deception to some extent using the tools we already have. Some intrusion prevention appliances implement deceptive measures at the network protocol layer. TCP tarpits, too, can be configured to respond to TCP handshakes without opening a real connection; a few EDR (End point detection and response) tools are capable of allowing deception at the malware host layer. While certain attack vectors can be captured with these methods, they also have their own limitations. Modern HoneyNet (MHN) is a popular open source honeynet. There are several vendors operating in this space, including:

Table 2:

Company	Year
Rapid7	2000
LogRhythm	2003
ForeScout	2000
Shape Security	2011
Attivo Networks	2015
Acalvio Technologies	2015

Illusive networks	2014
GuardiCore	2013
Cymmetria	2014
TrapX Security	2012
Minerva Labs	2014
CounterCraft	2015
Allure Security Technology	2011
CyberTrap Software	2015
Smokescreen Technologies	2015
Hexis Cyber Solutions	2013
TopSpin Security	2013
Ridgeback Network Defense	2014
Percipient Networks	2014
Symantec Endpoint Protection	2017

Source: https://blog.aimultiple.com/deception-tech-companies/

4 Application Whitelisting

Application Whitelisting (AWL) is a realistic practice of ensuring only permitted programs and software libraries to be present and active, while others are prevented from execution. It mainly focuses on protecting computers and networks from potentially harmful applications. In direct opposition to the concept of Blacklisting, Application Whitelisting is a more proactive approach that allows only pre-approved and specified programs to run.

Application Whitelisting has the ability to manage, reduce or control the demand on resources within a network. The AWL products have the capability to prevent zero day attacks, which is an advantage of AWL offers to financial sectors, as it allows only a set of trusted applications and blocks everything else.

How application whitelisting works?

Each executable in the whitelist is uniquely identified typically by File name, File size, File path, Digital signature and Hash. Apart from these attributes, it also analyses the types of application resources

handled, like executables, libraries, scripts, etc., for generating a whitelist.

Things to consider while implementing AWL:

- Identify the operating system, applications, and the interfaces of other integrated systems, etc. that needs to be monitored
- Decide whether to “block execution” or “ just monitor” when unauthorized applications run
- Know in advance about the systems with built-in whitelisting, and the systems that require a product to be purchased (like Windows has a built-in product)
- Scanning of the system for malware is mandatory, before installing AWL
- Test and ensure that AWL does not block much needed security updates to run
- By executing few applications, test whether the AWL technology is blocking the applications that are not whitelisted and executing those that are in whitelist. Ensure that it is not doing vice-versa
- Deploy personnel dedicated to monitor the system.

4.1 Management of AWL

Implementation of application whitelisting is quite simple. It begins with building a list of approved applications. The whitelist can be built into the host operating system. A program that has to be executed is compared against a whitelist, and accordingly is permitted or blocked.

Once the system starts functioning, an IT admin must ensure that it is up-to-date, with latest whitelists and apply patches to the programs, increase the scope of AWL to other platforms and perform periodic tests to ensure the software is working as expected. The admin must go through the log file checking

particularly for the events of application blocks that might have blocked genuine whitelisted applications.

Advantages

- AWL places control over programs that can run on a computer system, in the hands of the administrators rather than the end users. The users cannot run malicious programs even inadvertently as the programs are not whitelisted by the administrator
- As AWL allows a set of trusted applications and blocks everything else, it provides more security from unknown malicious programs. This would even stop dreaded "zero day" attacks which is impossible with standard blacklisting
- The volume of malware is growing day-by-day. It is impossible to maintain a comprehensive blacklist. In today's environment, reverse approach is much more practical.
- Apart from permitting accepted applications to run on a system, AWL has the following capabilities as well:
  - It can provide file integrity i.e. changes to files can be stopped or monitored or reported
  - It can act as an asset database i.e. it can keep an inventory of applications and their versions that are installed on hosts
  - It keeps track of certain features of files such as cryptographic file hashes, sizes, date created, etc., which can be used to detect and respond in case of a malicious activity.

Disadvantages

- Whitelisting is not a replacement to traditional security measures. It is a supplement. AWL in combination with standard security technologies ensures enough protection
- Users who install a safe application for legitimate reasons face some inconvenience as they need to ensure that the application is whitelisted by the administrator

- When the organisation relies on a deny-by-default mechanism, a user must have an application whitelisted before s/he is able to use it. This process creates workflow delays that in turn lead to frustration in employees.

4.2 AWL Solutions

Airlock Digital; Lumension; Carbon Black; McAfee Application Control; Digital Guardian, etc., are some of the commercial offerings.

The free software like Applocker is included with Microsoft 7, 8, and 10; Gatekeeper is Apple's whitelisting solution; SELinex is Linux's whitelisting application control

Precautions

- Make sure the computer is clean before installing any solution. If malware is already present, the solution will readily whitelist it and allow it to run
- Routine maintenance of whitelist needs to be done on priority
- Make sure the solution covers both executables and software libraries. An omission of either can compromise whitelisting security
- Whitelisted executables must be identified by other means rather than merely filename or directory location so that the malware cannot trivially masquerade as legitimate software.

5 Continuous Security Validation (CSV)

Application Whitelisting (AWL) is a realistic practice of ensuring only permitted programs and software libraries to be present and active, while others are prevented from execution. It mainly focuses on protecting computers and networks from potentially harmful applications. In direct opposition to the concept of Blacklisting, Application Whitelisting is a more proactive approach that allows only pre-approved and specified programs to run.

Continuous security validation promises to keep the infrastructure safe and secure; and helps organisations address unmitigated risks created by today's ever-increasing threats. Despite placing enough security controls, deploying security solutions and employing good security practices, organisations are not sure and confident that the mechanisms in place are efficient. In the world of security, one thing is sure – that the unknown can hurt badly.

The variety of security products with gaps in their coverage, their misconfigurations, the visibility they offer always lag behind the new exploits. This kind of situation leaves behind residual risks for an organisation; leading to major security incidents.

Continuous security validation is like carrying a penetration testing with latest threat intelligence. A continuous security validation, allows security teams to:

- Determine whether security architecture provides enough protection
- Ensure security configurations are properly implemented
- Continuously assess vulnerabilities against active threats.

Problems with penetration testing and red teaming processes:

- Testing frequency: However frequent testing one may do, it falls short in the present dynamic IT environment
- Testing methodologies. Many testers repeat the same playbook for every test for months and years without updating their skills
- The highly technical output from VAPT exercises makes it complex for business teams to assess risk.

IS Audit and new needs

When it comes to security, auditors have the task to perform controls validation. Simply verifying that controls are in place, makes no sense, rather validating that those controls are effective is what matters:

- Shifting from reviewing configurations to validating configurations work

- Moving from a point-in-time to a continuous audit
- Transitioning to automation for multiple mandates
- Generating evidence-based security trends – not assumptions
- Operating with less dependence on other individuals and groups.

How continuous security validation works

These solutions are motivated from sayings like “Think like a hacker to break and put controls safeguard”. “It’s time to play the hacker. Complement existing reactive and defensive security with offensive tactics. Think and act like a hacker to gain a better understanding of how exactly you will be breached.”

The solution simulates hacker breach method and its ultimate impact. This allows teams to validate security defense controls, challenge SOC teams and more importantly disrupt and disable critical attack paths by applying fixes/patches.

The platform comes with a threat intelligence capability; and preloaded with lot of breach methods; malicious URL and IP address information etc., and allows creation of unique virtual customer environments, which are attacked by various exploits. A mix and match of security controls and systems can be simulated to validate how threats operate within a given configured environment and determine which risks remain open by that particular combination of controls.

CSV tools use a variety of current hacker techniques to hammer corporate networks continually. These tools then gather this data in dashboards that can be used to constantly monitor vulnerabilities and track remediation progress. Some tools can align results with the MITRE ATT&CK framework, and the best tools provide role-based dashboards for positions like security analysts, IT auditors, and business managers so they can review the results and use real-time data to prioritize remediation decisions. CASV tools and services come from vendors like AttackIQ, Censys, Metronome, Pycsys, Qadium, Randori, SafeBreach, SCYTHE, Verodin, CAWS from NSSlabs and others.

Best Practices for implementing CSV

To implement and execute Continuous Security Validation, a company could leverage industry best practices. A leading framework in this area is MITRE ATT&CK™ for Enterprise (ATT&CK). ATT&CK for Enterprise is a framework that takes the perspective of an adversary trying to hack into a company using various known attack vectors. This framework provides a library of real-world hacking activities for companies to simulate in their own networking environment.

In its simplest form, an organization could pick a relevant attack vector (e.g. exfiltration over alternative protocol) from the ATT&CK Matrix and test its cyber defenses to validate that it could withstand that particular attack. They can then review and prioritize mitigation of identified gaps.

6 Zero Trust Networks

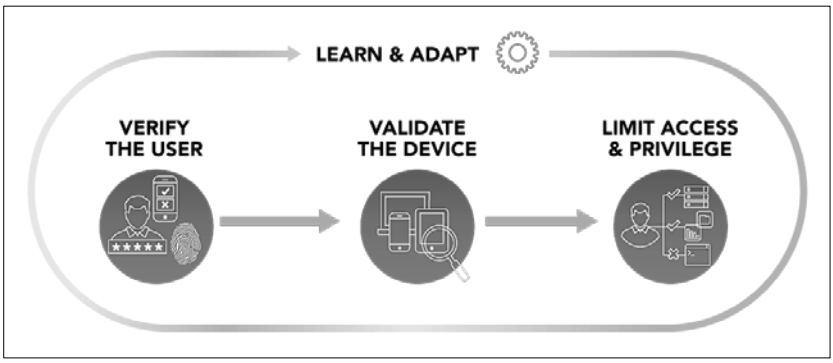
Zero trust networks is basically built upon five fundamental beliefs:

1. The network is always considered to be hostile.
2. It is considered that internal and external threats always exist in the network
3. The locality of the network alone is not enough to decide trust of a network.
4. Within the network every user, device and network/data flow must be authenticated and authorized.
5. The polices must be made considering various factors and they must be highly dynamic as well.

Usually security models operate on the assumption that everything inside the organization's network can be trusted on, but attacks also happen due to insiders knowingly or unknowingly. The zero trust model works on the principle: never trust, always verify. The

architecture is such that it verifies everything and everyone without trusting anyone.

6.1 Zero trust model



Source: https://blog.centrify.com/best-practices-zero-trust-security/

The zero trust models work on the principle that it moves access control mechanisms from perimeter of network to actual devices, systems and users. A centralized policy (policy engine) verifies the identity of user and validates his device before granting access permissions, but it smartly limits the access permissions and privileges one can have on the target system.

Apart from the centralized policies, dynamic variables like the previous behaviour of the user, location of the user, day and time to form the trust score for each access attempt made are important. Based on this trust score, the access is granted or denied to the system.

The following are the three operational objectives that define zero trust implementation:

Objective 1: All the resources must be accessed securely regardless of the location. Multiple trust boundaries must be formed for communication from and to the resources, even when the communications are confined to the internal network. Only devices with right settings and status must be allowed to access the system/network.

Objective 2: Strict access control and the least-privilege policy must be followed. In this the goal is to reduce the allowed access to resources as a remedy to reduce the possibility for attackers and malware to gain unauthorized access, and get into critical data.

Objective 3: All traffic must be inspected and logged. Strict enforcement of access control alone is not sufficient. Close and continuous attention must also be paid to exactly what is happening in "Allowed" application.

6.2 Zero trust best practices

- All the resources and data must always be accessed securely. The user location must also be logged to ensure security. It is very important to understand who the users are and what applications they are using. The method of connection is the best way to enforce and determine policies that ensure 100% security
- Access control must be enforced in the organization and a "least-privilege" access must be put
- All the traffic and logs must always be verified
- Add more authentication methods to counter credential based attacks
- Authentication rules must be made even stricter for credential based attacks
- Always keep adding new context to the organizations security and one golden rule is: never trust.

Benefits of adopting zero trust practices and principles in financial sector

The financial organizations can rely on Zero Trust architecture for preventing the exfiltration of sensitive data and improving the defence mechanism against changing cyber threats.

- Using zero trust solutions makes non-disruptive and incremental transition to the zero trust model

- Establishing a zero trust network gives us awareness about enterprise computing and security activities
- Policies like least privilege access and other concepts can easily be implemented which reduces the attack surface
- Implementing zero trust network will improve the company's security posture and ability to protect critical data from leakage
- Zero trust networks can achieve compliance with applicable standards and regulations, simply by using good trust boundaries
- As financial organizations are business-driven, using zero trust networks will help in initiatives like cloud computing, user mobility, virtualization and social networking
- Since zero trust networks use single consolidated platforms for security across the entire computing network, instead of disconnected points of products at various levels, it reduces the total ownership cost.

7 Threat Intelligence

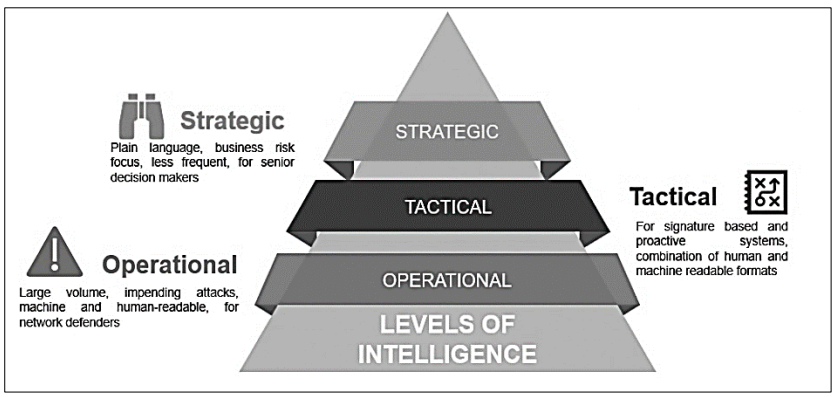
Traditionally, security teams use a variety of tools and techniques to conduct threat analysis, incident response and organization network defense. Information sharing between the security team and other teams is usually a manual or a ticketing system. Changing attacking methodologies demand upgrading the knowledge of the security team. Threat intelligence fastens this process and benefits the organization with latest threat information. Threat intelligence, also called as cyber threat intelligence, is the analyzed (both internal and external threats), refined and organized information about potential or current attacks, that threaten an organization. The threat intelligence can be gathered against many threats, like:

- **Zero day threats:** It is a computer-software vulnerability, which is unaddressed by the ones responsible for addressing the (patch management team, security testers). From the name of the threat, it is clear that software is vulnerable since the day of release.

- **Exploits:** Once the attacker finds a way into the network, his motive is to take control of the entire network/system and exploit it.
- **APT (Advanced Persistent Threats):** Sometimes the attacker's motive might not be to destroy the network/system; his aim would be to remain in the network unnoticed and rob all the possible information.

All such kind of threats can be detected by using threat intelligence in the right manner.

7.1 Levels of threat intelligence



Source: https://www.crest-approved.org/wp-content/uploads/CREST-Cyber-Threat-Intelligence.pdf

- **Strategic Intelligence:** This intelligence gives us the "who and why": who is attacking the organization and what is their intention. It is an eagle eye view of the current attacks and threats. The decision makers of an organization (board members, senior management and executives) use this intelligence, to bring in required resources with the knowledge gained from the strategic intelligence.
- **Tactical Intelligence:** Tactical intelligence will help defenders (SOC analysts, patch and vulnerability management teams,

incident response members) understand what has happened to prevent further attacks on the victim organization, which has recently been attacked. This intelligence is a ground level view, it describes about IOC (indicators of compromise) which are associated with known attacks.

- **Operational Intelligence:** This intelligence gives us an insight into the TTP's (tactics, techniques and procedures) like their phishing schemes, new tools, IP addresses, protocols used by the attackers. This knowledge helps us successfully to establish defense throughout organizations. The operational intelligence is of use to the crisis management and IT operations.

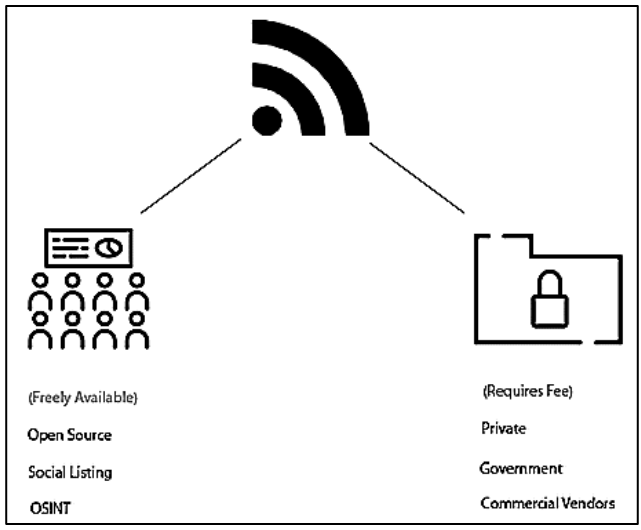
7.2 Threat intelligence feeds (TI feeds)

As the name suggests, they are threat intelligence data/IOC that can be fed to the security devices like Security information and event management (SIEM).

For these feeds, the organization must assess itself, i.e., depending on the organization's domain. There are different kinds of feeds and they can be combined based on the purpose and requirement. Threat data is gathered from human intelligence, signal intelligence, open source intelligence, geospatial intelligence, financial intelligence, market intelligence and tech intelligence in real-time.

The TI feeds are categorized broadly into two types:

- **Free or publicly available feeds:** These are available on the internet and is also called as OSNIT (open source intelligence)
- **Private feeds:** They need to be purchased from security vendors depending on the requirement.



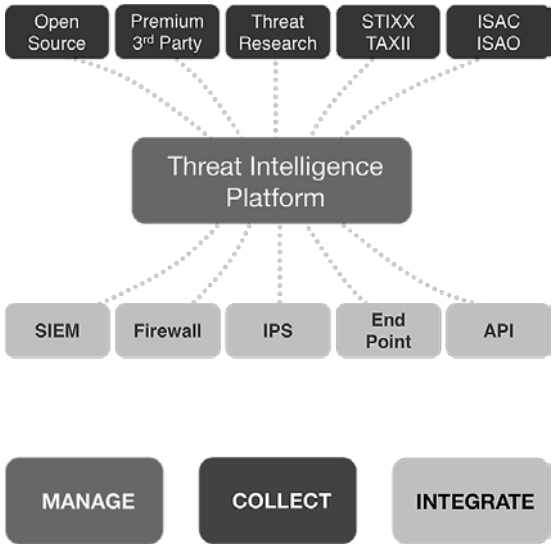
Source: https://www.cm-alliance.com/cybersecurity-blog/importance-of-threat-intelligence-feeds

Table 3:

Public Threat Intelligence	Private Threat Intelligence
• Open Source Feeds • Social Media • Pastebin • Using Trusted Automated eXchange of Indicator Information (TAXII) • Commercial • Government reports • Internal Sensors	• SHODAN • ThreatConnect • Virus Total • AlienVaults OTX (open threat exchange) • Zeus Tracker • The dark web from where you can obtain feeds

7.3 Threat intelligence platform (TIP)

It is a platform used to help the organization aggregate, correlate, remove de-duplicated data and analyze threat data from multiple sources(open source and paid) in real time.



Source: https://www.anomali.com/resources/what-is-a-tip

TIP capabilities

- **Collect:** Collect data from various sources
- **Correlate:** Know the relationship between different threats if any – correlate
- **Enrichment and contextualization:** Get threat data from other organizations of the same domain and enrich your data
- **Analyze:** Analyze all the collected threats
- **Integrate:** The collected, correlated, analyzed data must be, integrated with existing security systems.

Advantages of using TIP:

- The first and principal benefit of real time cyber threat intelligence is that it is available in real-time
- Access to information on new and emerging threats and threat actors
- The ability to track the ongoing activities of cyber-criminals and attacks specific to your industry.

Threat Intelligence Solutions
IBM X-Force Exchange
Anomali ThreatStream
Palo Alto Networks AutoFocus
Palo Alto MineMeld RSA NetWitness Suite
LogRhythm Threat Lifecycle Management (TLM) Platform
FireEye iSIGHT Threat Intelligence
LookingGlass Cyber Solutions
AlienVault Unified Security Management (USM)

8 Conclusion

In this paper, we tried to provide an overview on six emerging solutions for cyber security, namely: 1. DLP-Data Leakage Prevention, 2. Deception Technologies, 3. AWL – Application Whitelisting, 4. Continuous Security Validation, 5. Zero Networks; and 6. Threat Intelligence Solutions.

References

1. https://www.ey.com/Publication/vwLUAssets/EY_Data_Loss_Prevention/$FILE/EY_Data_Loss_Prevention.pdf
2. https://www.ten-inc.com/presentations/2017_ISE_NA_NSSLabs_WP.pdf
3. https://aspioneer.com/safebreach-a-pioneer-in-the-breach-and-attack-simulation-market/
4. https://www.cybonet.com/images/CyBoWall/Network_Traps_Whitepaper.pdf
5. https://www.csoonline.com/article/3387616/the-case-for-continuous-automated-security-validation.html

6. http://www.isaca.org/Knowledge-Center/Blog/Lists/Posts/Post.aspx?List=ef7cbc6d%2D9997%2D4b62%2D96a4%2Da36fb7e171af&ID=1219
7. https://www.paloaltonetworks.com/cyberpedia/what-is-a-zero-trust-architecture
8. https://www.centrify.com/education/what-is-zero-trust/
9. https://www.csoonline.com/article/3287057/network-security/what-it-takes-to-build-a-zero-trust-network.html
10. https://www.oreilly.com/library/view/zero-trust-networks/9781491962183/ch01.html
11. https://www.forrester.com/report/Five+Steps+To+A+Zero+Trust+Network/-/E-RES120510
12. https://www.esecurityplanet.com/network-security/deception-technology.html
13. https://www.forbes.com/sites/danwoods/2018/06/22/how-deception-technology-gives-you-the-upper-hand-in-cybersecurity/#161d082a689e
14. https://earlyadopter.com/2018/06/13/active-defense-how-deception-has-changed-cybersecurity/
15. https://www.networkworld.com/article/3019760/network-security/the-ins-and-outs-of-deception-for-cyber-security.html
16. https://www.thewindowsclub.com/what-are-honeypots
17. https://en.wikipedia.org/wiki/Honeypot_(computing)
18. https://searchsecurity.techtarget.com/definition/honey-pot
19. https://www.techopedia.com/definition/10278/honeypot
20. https://www.titanhq.com/blog/how-do-you-implement-honeypots-in-your-organization
21. https://github.com/paralax/awesome-honeypots/blob/master/README.md
22. https://pdfs.semanticscholar.org/presentation/d6a3/ad5c2cf61cbf6d0edfb2eb92a44198baa577.pdf
23. https://focus.forsythe.com/articles/509/Deploying-Data-Loss-Prevention-Best-Practices-for-Success
24. https://www.gb-advisors.com/data-leak-prevention-dlp/
25. https://www.bankinfosecurity.com/webinars/preventing-unauthorized-access-to-your-institutions-data-w-119
26. http://www.cbai.com/news/What%20is%20DLP%20and%20Why%20is%20it%20Important%20to%20My%20Bank.pdf

27. https://www.bankinfosecurity.com/webinars/preventing-unauthorized-access-to-your-institutions-data-w-119
28. https://www.gb-advisors.com/data-leak-prevention-dlp/
29. https://www.veracode.com/security/data-leak-protection
30. https://www.esecurityplanet.com/network-security/data-loss-prevention-dlp.html
31. https://ostec.blog/en/perimeter/dlp-what-is-it-and-how-does-it-work
32. https://spinbackup.com/blog/data-loss-prevention-tool-advantages/
33. https://securosis.com/assets/library/reports/DLP-Whitepaper.pdf
34. https://resources.infosecinstitute.com/data-loss-prevention-dlp-strategy-guide/#gref
35. https://nvlpubs.nist.gov/nistpubs/specialpublications/nist.sp.800-167.pdf
36. https://www.calyptix.com/top-threats/application-whitelisting-good/
37. https://searchsecurity.techtarget.com/definition/application-whitelisting
38. https://www.linkedin.com/pulse/tiered-communications-model-cyber-threat-intelligence-shane-anglin
39. https://securityintelligence.com/what-are-the-different-types-of-cyberthreat-intelligence/
40. https://www.cm-alliance.com/cybersecurity-blog/importance-of-threat-intelligence-feeds
41. https://threatconnect.com/blog/strategic-vs-tactical-threat-intelligence/
42. https://www.anomali.com/blog/what-is-tactical-threat-intelligence.

* The contribution made by Ms. Lahiri Kolli, Research Associate, IDRBT is gratefully acknowledged.

Financial Inclusion: Emerging Role of FinTech

– **Dr. M. V. N. K. Prasad,**
Associate Professor, IDRBT

Abstract

Financial inclusion plays an indispensable role in inclusive growth of an economy by addressing the challenge of poor access of financial services to rural masses. Through this paper, an attempt has been made to provide an overview on the status of financial inclusion and the role of FinTech in past few years.

1. Literature Review

Financial inclusion is an essential keystone for economic development. A survey indicates that for enhancing the economic status of the financially feeble, a push towards financial inclusion is a critical step. It comprises of a collection of activities such as insurance and savings and is not restricted to expansion of credit facilities [1]. In 2005, experts observed that financial services could garner profits with 'low margin-high volume' by inclusion of large number of people at the bottom of pyramid. Therefore, banks need to relook at their business plans to make financial inclusion more possible for low income groups. In addition, technology and other resources must be used in the ways possible to accomplish the primary motto of financial inclusion [2].

In rural areas of Vellore district, Tamil Nadu, a survey was conducted in 2008 to examine and determine distinct motivating factors for financial inclusion. By accumulating data from 20 villages of Vellore, the researchers examined the reasons behind less number of accounts. The main reasons for lack of financial access were identified to be poor literacy rates/higher illiteracy, unemployment and low income levels [3]. Another study in Gulbarga district of Karnataka found that financial inclusion was up to 100%, especially due to the opening of National Rural Employment Guarantee Programme (NREGP) scheme bank accounts. The survey figured the necessity to increase financial literacy among rural people and create awareness about financial inclusion [4].

2. What is Financial Inclusion

The recent years have seen numerous reports on financial inclusion. Financial inclusion, as a niche area in financial sector, has gained substantial consideration of start-ups, regulators, venture capitalists and traditional financial institutions among developed and developing countries. Financial inclusion may be elucidated as having inadequate access to and utilizing financial services, which meet the user's requirements like payments, savings, credit and insurance conveyed in an accountable and feasible way. Financial inclusion can also be termed as a key enabler for two reasons: i) extreme poverty reduction for facilitating economic growth; and ii) access to financial services can act as a catalyst for digital economy.

The first step for extensive financial inclusion is having access to a transaction account through which people accumulate money, send and receive payments. Financial inclusion helps in day-to-day existence and guides the poor households to plan for long-term goals as well as emergencies. Accountholders are more likely to utilize additional financial services, start and extend businesses, contribute to health or education, etc., all of which enhances the overall nature of their lives.

According to recent Findex information, only 1.2 billion people have an account and 1.7 billion people approximately, all over the world, do not hold a bank account. Universally, two-thirds of people don't hold a bank account due to shortage of money, lack of proximity with financial service providers, documentation and trust – which indicates that more efforts need to be channelized to devise tools to provide financial services for low-income users.

Over 55 countries since 2010 have committed to financial inclusion and over 60 countries have either started or are preparing a national strategy. When countries adopt a strategic method and establish financial inclusion approaches, financial regulators, competitors,

telecommunications and education ministries join in. Countries that have accomplished the development towards financial inclusion have:

- Leveraged government payments (For instance, 35% of adults in low income countries accept government money by opening bank account)
- Mobile financial services are allowed to bloom. (For instance, mobile money account ownership increased from 12% to 21% in Sub-Saharan Africa.)
- Novel business models were encouraged, like extracting e-commerce information for financial inclusion

3. Evolution of FinTech

The year 2015 was a big year for the Indian FinTech sector, for it saw the growth of various incubators, FinTech start-ups, as well as from private and public investments.

An appropriate mix of capital investments, technical skills, regulatory framework, government policies, and creative and entrepreneurial mindset would be the driving forces for FinTech to keep picking pace. Framing a booming FinTech ecosystem where start-up firms partner with universities, research and financial institutions, technology experts and government agencies, is also going to be vital step for development and modernisation of the FinTech sector.

4. Status of Financial Inclusion in India

Up to 2017

Policy makers and researchers considered financial inclusion as a crucial enabler to accomplish equitable growth. Across the world as well as in India, policy makers made efforts to make Financial Inclusion a game changer, by bringing in the unbanked and needy to become a part of the financial system. To achieve complete Financial Inclusion, evaluation of the current status of financial inclusion is essential.

For this purpose, distinct exploratory studies and analysis have been carried out. To assess the scope of financial inclusion, some researchers computed Financial Inclusion indices. Dr. K. C. Chakrabarty, former Deputy Governor, RBI [5] had stated that almost half the country is unbanked and financially excluded households then accounted for 14.50 Cr. Among all, only 55% had deposit accounts and the percentage of current accounts was 9%.

CRISIL (2009), a credit rating agency, reported that the number of financially excluded households in India were around 12 crore (120 million). Rangarajan Committee (2009) [6] inferred to NSSO information reporting the scope of Financial Exclusion and stated that out of 89.3 million households, 45.9 million farmer households in the country (51.4%) don't have access to credit, either from non-institutional or institutional sources. It also mentioned that regardless of the extensive network of bank branches, out of the total farm households, only 27% are indebted to formal sources (wherein one-third of them also borrow from informal sources). Farm households not having access to credit from formal sources as a proportion to total farm households is specifically high at 95.91%, 81.26% and 77.59% in the North Eastern, Eastern and Central Regions respectively. Hence, apart from the evidence that exclusion in specific is large, it also alters broadly across regions, asset holdings and social groups. It concluded that exclusion was high among financially weaker groups.

Shri H. R. Khan, former Deputy Governor, RBI [7], stated that along with other states in India, UP has very low Financial Inclusion. He suggested three dimensions to find the scope of Financial Inclusion viz. scope of banking services, banking penetration and utilisation of banking services. As per research work in (2008) [7], the author developed a multidimensional index (Index of Financial Inclusion, IFI) that collects data on distinct dimensions of Financial Inclusion in a single digit that varies between 0 and 1, where 0 signifies complete Financial Exclusion and 1 denotes complete Financial Inclusion in the

economy. By utilizing the three dimensions, the author computed IFI for the year 2004 for 100 countries wherein India holds 29th rank. By applying the same technique for the districts of Punjab, in (2011) [8] the author stated that financial inclusion was very low with 3d-IFI value at 0.354, whereas, financial inclusion was high in six districts namely SAS Nagar, Kapurthala, Jalandhar, Patiala, Ludhiana and SBS Nagar and low in the other 11 districts out of 20.

In a research paper [9], the authors mentioned that financial inclusion is similar to banking inclusion and also suggested an index that was deployed in building human development index by utilizing axiomatic measurement scheme. As per the suggested scheme, in the overall accomplishment of financial inclusion, the individual dimension's contributions percentage can be calculated by using the proposed index of financial inclusion to recognize the dimensions of financial inclusion, which are more or less accountable for overall inclusion.

To evaluate the scope of financial inclusion in India, CRISIL recently came up with a composite index namely CRISIL Inclusix. This index is based on three drivers/dimensions of financial inclusion namely credit penetration, branch penetration and deposit penetration by banks, which is used as a comprehensive competitor to measure the scope of financial inclusion at district, state, regional and national level. In this index, four divisions are ranked to illustrate distinct levels of financial inclusion as follows:

Four categories for CRISIL Inclusix	
CRISIL Inclusix Score	Level of Financial Inclusion
>55	HIGH
Between 40.1 and 55	ABOVE AVERAGE
Between 25 and 40	BELOW AVERAGE
<25	LOW

Table 1: Categories for CRISIL Inclusix

CRISIL released three volumes in June 2013, June 2014 and June 2015 to report the scope of financial inclusion based on the four above-mentioned categories. In the first volume, CRISIL Inclusix score has been increased by the reason of enhancement in deposit penetration. For the overall improvement in CRISIL Inclusix score, the authorities focussed on the other two parameters - credit and branch penetration. The CRISIL Inclusix score for the whole country was 40.1 and for UP the score was between 25 and 40, which clearly shows that it falls under below average level of financial inclusion **(Table 1)**. In the second volume, CRISIL Inclusix score for the country was 42.5 and the score for UP increased to 35.2. Lastly, in the third volume, microfinance institutions contributed for the first time in the computation of CRISIL Inclusix score and the score for the country was 50.1. The increased level of financial inclusion was mainly because of improvement in UP score to 40.1. India ranks low when compared to other countries in financial inclusion and within India, UP is lagging behind when compared to other states.

2018

Experimental survey has proposed that specific indicators are needed for developing financial inclusion polices. Some effective indicators of financial inclusion are defined by the World Bank, International Monetary Fund and other global organizations. A few of these key indicators are the number of ATMs installed, number of bank branches, deposits, bank credit and so on. As per the research, when measured per 1000 km of region, the most populated country i.e., China, has a broad network of financial inclusion having 1428.98 bank branches.

In India, to enhance the scope of financial inclusion, the government of India has suggested all private and public sector banks to devise a financial inclusion plan (FIP) for three years, which will include data about the prominence of Kisan Credit Cards (KCC) in rural areas, the number of brick and mortar bank branches, number of General Credit Cards (GCC) and so on. Table 2 presents a snapshot of growth of

financial inclusion in India. In a country, banks are the key providers of financial access and one of the strong indicators of the scope of financial inclusion is the number of bank branches. The increase of bank branches from 2013 to 2017 are shown in **Table 2**.

S. No.	Particulars	March 2010	March 2011	March 2012	March 2013	March 2014	March 2016	March 2017
1	Banking outlets in Rural locations – Branches	33378	34811	37471	40837	46126	51830	50860
2	Banking outlets in Rural locations – Branchless mode	34316	81397	144282	227617	337678	534477	547233
3	Banking outlets in Rural locations – Total	67694	116208	181753	268454	383804	586307	598093
4	Urban locations covered through BC*	447	3771	5891	27143	60730	102552	102865
5	Total Kisan Credit Cards (KCC, No. in million)	24.3	27	30	34	40	47.3	46
6	KCCs – Total (Amount in Rs. billion)	1,240	1,600	2,068	2,623	3684	5,131	5805
7	Total General Credit Cards (GCC, No. in million)	1,4	2	2	4	7	11.3	13
8	GCC – Total (Amount in Rs. billion)	35	35	42	76	1097	1,493	2117
*BC: Business Correspondents								
Source: Compiled from Report on Trend and Progress of Banking in India of various years (issued by RBI)								

Table 2: Financial inclusion plans for scheduled commercial banks along with RRBs

4.1. Reasons for FinTech Boom in India

In India, the growth of FinTech has been phenomenal, to the extent where the most widely used messaging app, WhatsApp, has introduced a new payment feature by using Unified Payments Interface (UPI). UPI, managed by the National Payments Corporation of India (NPCI), is an instant and real-time payment system between participating banks to transfer money through mobile-to-mobile. India accounts for over 200 million WhatsApp users providing a huge consumer base for merchant payments and contributes to large volumes in terms of peer-to-peer (P2P) payments.

Many other global IT giants are also zeroing in on adding payment feature. For example, in 2017, Google Tez, a payment app, was launched by Google, while Amazon has launched Amazon Pay and Samsung has introduced Samsung Pay. UPI feature has been enabled by Samsung Pay and Google Pay and in the near future, Apple plans to launch Apple Pay in the country.

In the meantime, UPI-enabled digital payment firms such as PhonePe, MobiKwik, Paytm and FreeCharge are beefing up their arsenal. In 2017, India's largest online payment firm and mobile wallet company, Paytm, invested Rs. 5,000 crores (786 million) in mobile payments. Flipkart, one of India's leading e-tailers, has also invested $500 million in PhonePe, to extend its technology platform and consumer base and scale up its merchant network through online payments. The CEO of PhonePe said that in the year 2017, PhonePe grew at over 100%, every two months.

In China, Tencent, a Chinese conglomerate owns the largest digital payments service i.e. WeChat. As per the Digital Payments 2020 report by Google-Boston Consulting Group (BCG), by 2020, India will exceed $500 billion in digital payments, up from $50 billion in 2016. BCG India managing director and senior partner said that "Globally, digital payments is sustaining with brisk transformation and by 2020,

it is set to increase four times in value". Analysts predict that by 2023, digital payments will exceed cash transactions.

5. Financial Inclusion and FinTech: Status all over the World

The progress of financial inclusion was a result of the thrust by governments, which introduced various schemes, digital payment modes using internet and mobile phones. Financial inclusion increased the access and utilization of accounts in Sub-Saharan Africa, where a mere 21% of people had accounts earlier. In East Africa, mobile money transactions have gained momentum; and it is picking up in West Africa as well.

Digital technology has altered the payments landscape. In 2017, the percentage of adults making digital payments was 52% as compared to 42% in 2014. Technology helped create customer awareness about access to financial services through digital mode. Using technology platforms, China also increased the number of account holders' usage for payment purposes, where 57% account holders are making purchases or paying bills through internet or the mobile phones – nearly twice the percentage in 2014.

Access of financial services to women have also improved. Three years ago, in India, the male account holders were more than female – a difference of 20 percentage points. Presently, this gap has decreased to 6 percentage points because of government's push.

However, in many countries, women continue to lag behind men in terms of account holding. Sixty percent of women hold accounts when compared to 72% men, globally, a gap unchanged since 2011.

Globally, many organizations pay wages to about 230 million unbanked adults. If these organisations decide to pay wages using mobile phones or the internet, it may help the workers become a part of the formal financial system. Private sector, development

organizations and the government need to work together to bring people into the formal financial system.

Universally, in 2017, about 1.7 billion adults neither had a mobile payment account nor a bank account (Figure 1) [10], show a slight decrease in the number of unbanked adults, as compared to 2 billion in 2014. Since account ownership is more or less universal in high-income economies, the unbanked adults exist in developing economies. Though account ownership share is high in China and India, share of unbanked adults is more in these countries because of their sheer size. The world's largest number of unbanked adults live in China, followed by India (190 million), Pakistan (100 million) and Indonesia (95 million) as depicted in **Figure 1**. In addition to these four economies, Nigeria, Mexico and Bangladesh are the other three countries, which constitute approximately half of the world's unbanked population **(Figure 2)**.

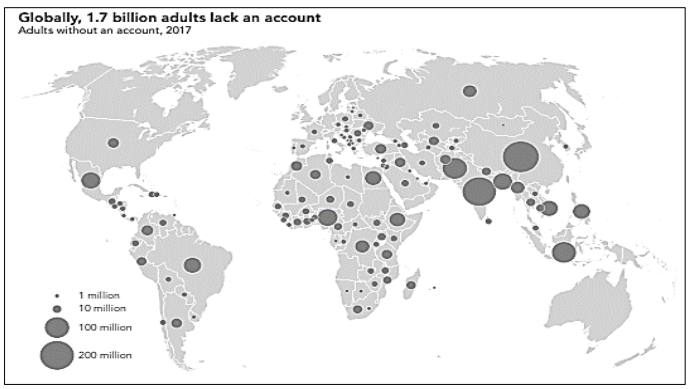


Fig. 1: Adults without an account, 2017

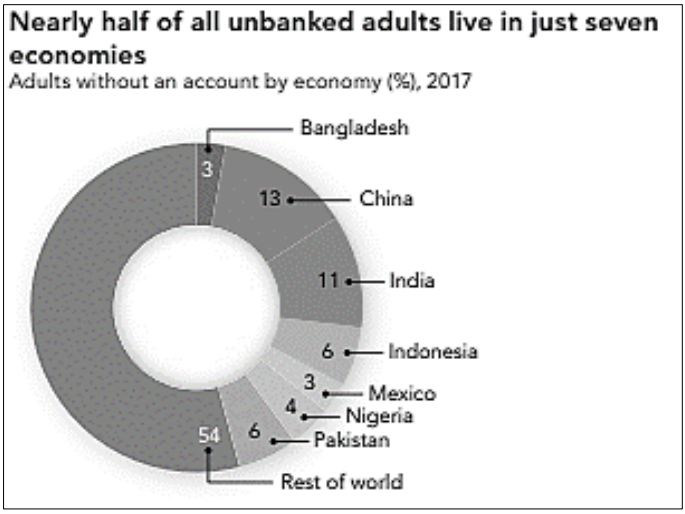


Fig. 2: Unbanked adults in seven economies (%), 2017

5.1. Reasons for unbanked adults

There are 1.7 billion unbanked adults all over the world i.e. they have no accounts either at banks or a mobile money provider. In developed economies, the number of bank account holders is higher as compared to the unbanked adults living in the developing world. Half of the unbanked population exists in developing economies like India, China, Nigeria, Indonesia, Bangladesh, Pakistan and Mexico and up to 56% of them are women. Adults without an account in 2017 by gender (%) is shown in **Figure 3**.

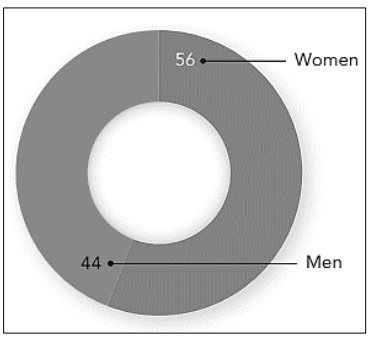


Fig. 3: Adults without an account by gender (%), 2017

Between 2011-2017, the share of account holders has increased globally, from 51% to 69%. Since 2011, the share of adult bank holders in India has doubled to 80%. The government's initiative for inclusive growth of banking for the unbanked in 2014 to boost account ownership through biometric identification cards was the key reason for the increase in numbers. Among women the share of account ownership is over 30% and among adults in the poorest households is 40%. In developing economies like China, Malaysia, Brazil and South Africa, the share of account ownership has been unchanged **(Figure 4)**.

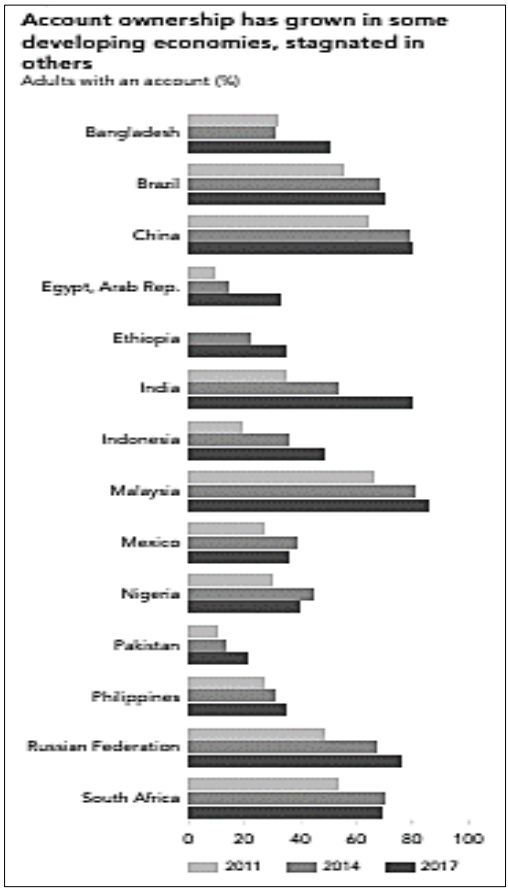


Fig. 4: The share of account owners in developing economies (%), 2017

As per 2017 Global Findex survey, it was reported that 76% of account owners (52% of adults) across the world received or made at least one digital payment **(Figure 5)**. In high-income economies, the share of account owners is 97% (91% of adults), while the share is 70% (44% of adults) in developing economies. The percentages include all people reportedly using either mobile phones or credit or debit cards to make payments from an account, or purchasing something online through internet or bill payments. The transactions also includes receiving or sending payments for agricultural products, paying bills or receiving wage, government transfers by a mobile money account, pension from or into a financial institution account.

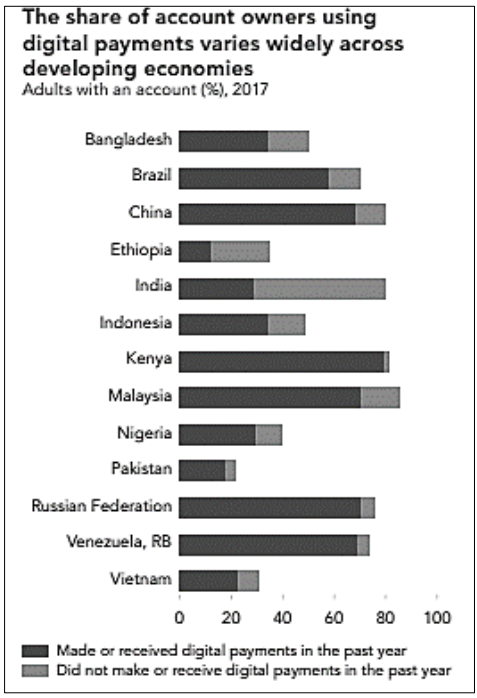


Fig. 5: Share of account owners using digital payments (%), 2017

Across the world, during 2014-2017, the share of adults with accounts rose by 11 percentage points from 41% to 52% either by receiving or making digital payments. In developing economies, the overall share of bank accounts rose by 12 percentage points due to use of digital

payments – among adults it rose from 32% to 44% and among account owners, the percentage rose from 57% to 70%. In Kenya, the percentage of people using digital payments is 97 and in China it is 85. Digital payments also helped increase the share of account owners in Venezuela and Russian Federation. Usage of digital payments is low in Ethiopia. The share almost doubled in Thailand to 62% and in economies like Malaysia, Brazil and Russia. (Figure 5).

To make payments, account owners use payment cards like credit or debit cards without having to withdraw cash. 80% of people use debit or credit cards in high-income economies to make payments while in developing countries only 22% do so. Debit card ownership is high in developing economies like Brazil, Malaysia, China, Turkey and Russia **(Figure 6)**. In Kenya and India, debit card usage is less than half of account owners and indeed only about a third of those who have used it to make a purchase. Only about a third use it to make a purchase in Indonesia, Egypt, Philippines and Nigeria. Debit card ownership is relatively high in Venezuela due to the country's economic challenges leading to shortage of bank notes. So, whenever possible, people use debit card to make purchases.

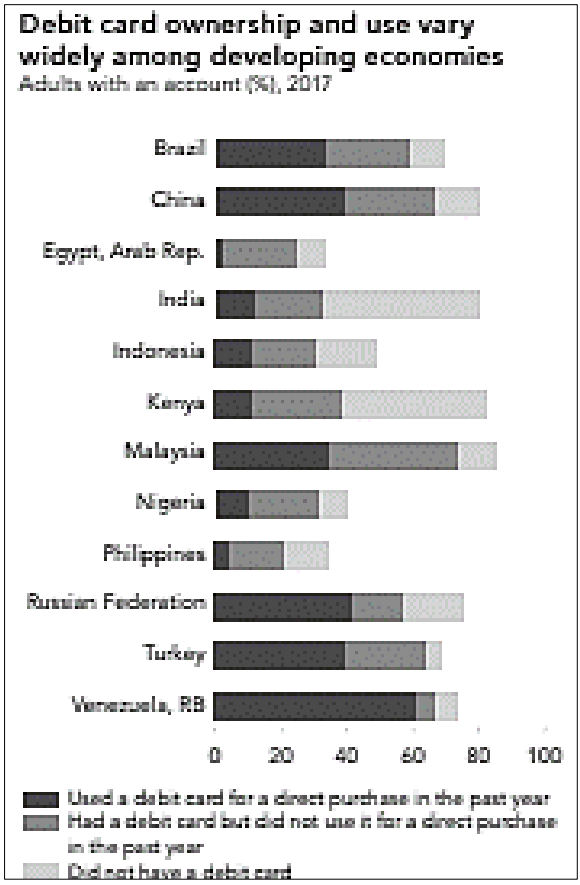


Fig. 6: The share of debit card ownership among developing economies (%), 2017

Throughout the world, millions of unbanked adults accept payments in the form of cash from the government, for the purpose of wages and sale of agricultural products. Account ownership increased by digitizing payments made through cash. Globally, in order to receive wages or government payments, 13% of account owners or 9% of adults have opened their first account **(Figure 7)**. In Malaysia, Zambia and the Islamic Republic of Iran, the share is higher for first time account opening – up to approximately 20% of account owners. And the share is 25% in Peru, Argentina, Turkey and the Russian Federation and about 40% in Kazakhstan and Arab Republic Egypt.

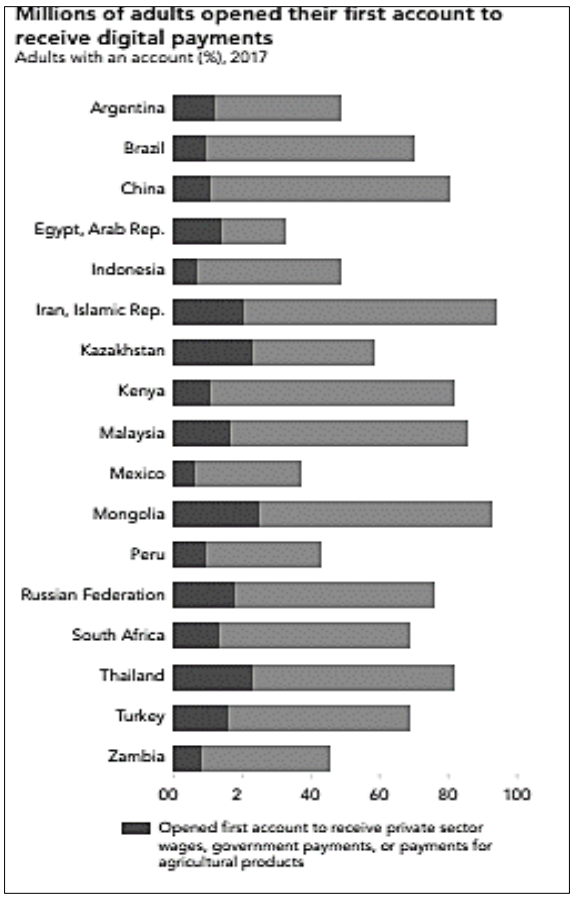


Fig. 7: The percentage of adults who opened first account to accept digital payments (%), 2017

The internet and the mobile phones are opening up new avenues to make transactions directly from an account either through a website or by an app or through using mobile money account. In the past year, financial transactions in high-income economies through internet or the mobile phones constituted 55% of account owners (51% of adults). The share is 85% in Norway, 33% in Japan, and 22% in Italy **(Figure 8)**. Whereas, in developing economies, 30% of account owners (19% of adults) performed such financial transactions. On the other hand, adults having mobile money account share is large in Tanzania

and Kenya. Kenya accounted for 70% of adults share (88% of account owners) and China had 40% of adults' share (49% of account owners).

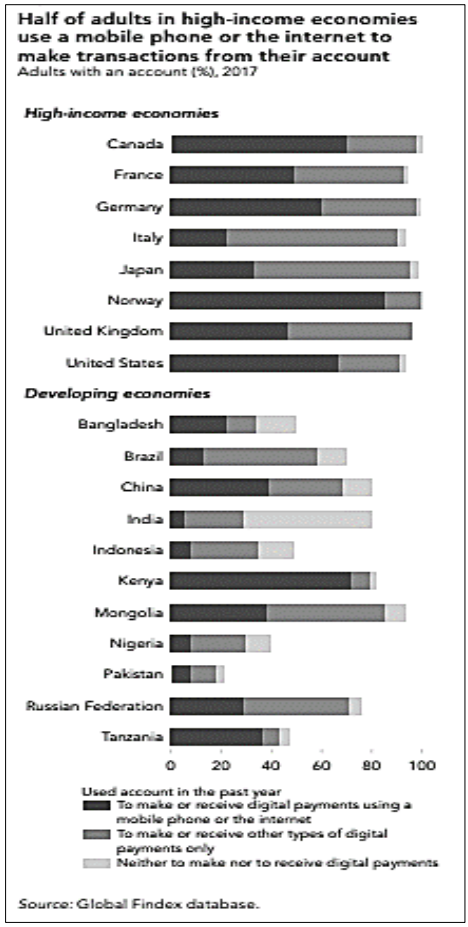


Fig. 8: The share of account owners using mobile phone or the internet (%), 2017

6. Privacy and security challenges for FinTech

The financial services sector manages sensitive information. With the evolution of FinTech, information is accessible in digital formats, and can be mined to obtain useful insights. This is also why the information is vulnerable to security attacks. According to global FinTech 2016

report of PwC, nearly 56% of the accused identified data privacy and security as hazards to the advancement of FinTech. Due to provision of online financial services and information ubiquity, information security has become a considerable challenge for FinTech. Since online services are increasing, tremendous amount of customer information is easily gathered, which can help determine acquisition and consumer buying patterns and retention schemes. Some of this information also contains personally traceable data and health and economic information. To secure this information and present it to third parties and consumers in a protected manner when needed are a challenge for the industry.

Alliance between new-age enterprises and traditional financial organizations have aided customers to have better goods and services at reasonable cost and enhanced access to existing goods and services – the key to such alliances is consistent information sharing. But consumer approval for information sharing requires stronger mechanisms. Care should be taken by organizations to create appropriate mechanisms to implement and reuse technologies as part of information lifecycle management and make sure that information is not misused. In order to safeguard data ownership, organizations need to standardize the systems both technically and legally. For instance, deleting consumer information when they unsubscribe utilization of FinTech services. Additionally, governing consumer access to services has also become complicated, which could be resolved by using cyber security tools.

The FinTech sector's major challenge is to protect the digital identities of people and organizations, by integrating multiple channels of services like banking and payments in a smooth way. The mobile phones enabled with biometric sensors (such as face scanner, fingerprint scanners) are being used for authorization and authentication instead of passwords and PINs. This could lead to more risks. Risk-based authentication or adaptive authentication could mitigate misuse of digital identities.

The enterprise applications, which communicate with interfacing systems made by application programming interfaces (APIs), shares the data but increases malware threat. Cross-platform malware contamination is risky. Viruses can produce malware, which can infect and proliferate from one platform to another. Withstanding such threat needs maintenance and regular updation of the systems and technologies used. By analyzing business use cases, embedding security, correlated controls and implementing threat schemes, assurance can be improved.

7. FinTech: Risks and Challenges for Indian Banks

The growing dependence of banks on IT has resulted in enhanced customer experience, reduction of response time and innovation of new technological products. At the same time, banks have witnessed decrease in operational complexities and other costs owing to outsourcing IT services to FinTech companies.

Few opine that outsourcing the data and other IT related services to FinTech companies poses information and security challenges. To mitigate such risks, it is essential for the banks to identify, assess and resolve the issues on a daily basis and fortify their risk culture for better governance and control.

As per a **BIS** (Bank for International Settlements) survey, majority of the FinTech service providers cater to services in the areas of payments, clearing and settlement, and capital-raising. There has been more increase in the number of retail payment services companies as compared to wholesale payment service providers. Banks need to chart out FinTech strategies keeping in view the risks associated with it.

Another concern for banks is the impact on their profitability due to unbundling services as customers explore multiple options of banking services, which are available at a lesser cost. In such a scenario, banks

must decide upon whether they compete or collaborate with the FinTech firms.

Outsourcing is another risk to be dealt with by the banks. As banks are outsourcing multiple IT services to various third parties, it is important for banks to put a strong system in place to watch over the operations and risks, enforce a strong contract and control assurance as a protective measure.

The third factor is the operational risk. According to BIS, operational risks which are emerging from FinTech, can be both distinctive and systematic in nature. The advancement of FinTech leads to more importance for technology, which may lead to an IT risk becoming a systematic crisis specifically where services are confined to one or few leading participants. Besides, banks should focus on developing new products and processes, alter governance processes to reduce the complexity of financial services delivery of these new services and products under operational risk. Moreover, legacy OS would be unable to cope up with the challenges arising from new services and products.

'Cloud sourcing' which is cost-effective enables banks to share essential resources (such as analytics and software packages), but leads to risks like 'data privacy, security, cybercrime, money laundering and consumer protection'.

Cyber risks may be on rise if business models and technologies do not keep pace with timely amendments. Because of reliance on APIs, other new technologies and cloud computing may aid and enhance interdependency with distinct FinTech institutions. But not complying with regulatory and cyber security norms will expose the banking system to cyber threats.

With the expansion of big data, the risk of non-adherence to privacy rules could increase due to outsourcing. The privacy rules varies from region to region leading to additional raise in compliance costs for

multinational banks. One of the fundamental elements of operational risk is related to people/staff which needs more attention. Training the staff and creating awareness about the systems and processes is the key to governing this risk at all stages.

8. Conclusion

However advanced be the technologies, the financial sector would not see expected progress until financial inclusion benefits the underdeveloped. FinTech should play a critical role in enabling one and all have access to the financial services which requires a 'sustainable, collaborative and healthy financial ecosystem'.

References

1. Dev, S. M. Financial inclusion: Issues and challenges. Economic and political weekly, pp. 4310-4313, 2006
2. Leeladhar V. Taking Banking Services to the Common Man – Financial Inclusion. Commemorative Lecture at the Fedbank Hormis Memorial Foundation at Ernakulam. 2005
3. Sriram, M., & Sundaram, N. Financial inclusion index: a customized regional model with reference to economically most backward districts of Tamil Nadu, India. Mediterranean Journal of Social Sciences, 6(6), 209, 2015
4. Ramji, M. Financial inclusion in Gulbarga: Finding usage in access. Institute for Financial Management and Research Centre for Micro Finance. Working paper, 2009
5. Khan, H. R. Issues and Challenges in Financial Inclusion: Policies, Partnerships, Processes & Products. Keynote address delivered at symposium organized by the IIPA, Bhubaneswar. https:// www.rbi.org.in/scripts/BS_SpeechesView.aspx?id=711, 2012
6. Rangarajan C. Report of the Committee on Financial Inclusion. http://www.nabard.org/reportcomfinancial.asp, 2008
7. Sarma, M. Index of Financial Inclusion, ICRIER Working paper No. 215, http://www.icrier.org/pdf/mandira, 2008
8. Kainth, G. S. Developing an Index of Financial Inclusion, http:// www.microfinancegateway.org/library/developing-index-financial-inclusion, 2011

9. Chakravarty S. R. and Pal R. Measuring Financial Inclusion: An Axiomatic Approach. IGIDR Mumbai. http://www.igidr.ac.in/pdf/publication/WP-2010-003.pdf, 2010
10. Demirguc-Kunt, Asli, LeoraKlapper, Dorothe Singer, Saniya Ansar, and Jake Hess. The Global Findex Database 2017: Measuring Financial Inclusion and the Fintech Revolution. The World Bank, 2018
11. Kuntal Sur. Top five risks posed by the fintech revolution to Indian banks. The Economic Times. April 2018.

Journal of Banking and Financial Technology

1. First Issue (in association with Springer) - January-June 2019
2. Third Issue - July-December 2018
3. Second Issue - January-June 2018
4. First Issue - July-December 2017

Access the Journal from www.idrbt.ac.in/jbft.html

IDRBT Staff Paper Series

1. Cyber Security - March 2019
2. Analytics - December 2017
3. Biometrics - October 2017
4. Cloud Computing - January 2017
5. Payment Systems - June 2016
6. Mobile Banking - September 2015

Access these Staff Papers from www.idrbt.ac.in/staffpapers.html

White Paper/Blueprint

1. White Paper on 5G Applications for Banking and Financial Sector in India
2. Blueprint of Blockchain Platform for Banking Sector and Beyond
3. White Paper on Applications of Blockchain Technology to Banking and Financial Sector in India

Access these from www.idrbt.ac.in/whitepapers.html

Frameworks

1. Cyber Insurance - A Reference Guide
2. Handbook on Information Security Operations Center
3. FAQs on Cloud Adoption for Indian Banks
4. Digital Banking Framework
5. Cyber Security Checklist
6. IT Vendor Management: Principles & Practices
7. Data Quality Framework
8. Cloud Security Framework
9. Green Banking Framework
10. Social Media Framework
11. Information Security Framework for Indian Banking Industry
12. Information Security Governance for the Indian Banking Sector
13. Holistic CRM and Analytics for Indian Banking Industry
14. Organizational Structure for IT in the Indian Banking Sector

Access all Frameworks from www.idrbt.ac.in/bestpractices.html

Published by:

IDRBT
Explore, Enable, Excel

Institute for Development and Research in Banking Technology
(Established by Reserve Bank of India)
Castle Hills, Road No. 1, Masab Tank, Hyderabad - 500057.
EPABX : +91 - 40 - 23294999, Fax : +91 - 40 - 23535157
Web : www.idrbt.ac.in E-mail : publisher@idrbt.ac.in