**International Journal of Data Informatics and Intelligent computing (IJDIIC)**
Vol.2, No.1, March 2023, pp. 51~61
DOI: 10.59461/ijdiic.v2i1.46

# Research on the analytics of traffic pumping in telecommunications via data science using rehabilitated frog leaping algorithm

**Imran Ullah Khan[1], Naga Lakshmi Sowjanya Cherukupalli[2]**
[1]Department of Electronics Engineering, Integral University, Lucknow, Uttar Pradesh, India
[2]CSE Department, Koneru Lakshmaiah Education Foundation, Green Fields, Vaddeswaram, Guntur, India

**Article Info**

***Article history:***

Received February 24, 2023
Revised March 18, 2023
Accepted March 21, 2023

***Keywords:***

Traffic pumping
Decision-support system
Cluster analysis
Decision trees
Rehabilitation frog leaping algorithm (RFLA)

**ABSTRACT**

One kind of telecom crime known as "traffic pumping" occurs when local phone companies artificially overstate the volume of calls flowing into their systems so that they may charge the calling party a greater access fee than their own. Lacking labels for training set makes it difficult to determine whether congestion pumps has occurred. In this study, we suggested a decision-support system based on cluster analysis and decision trees for identifying fraudulent cases. In this study, we use the IBM Telco and Cell2cell datasets. The gathered information can be preprocessed using normalization. When we have collected enough data, we use the rehabilitation frog leaping algorithm (RFLA) to divide up the possible incidents of fraud into distinct groups. Next, we used the cluster participation labels to build a decision tree that led us to the criteria that must be met in order to pursue legal action against the circumstances that raised red flags. Professionals in the field of telecommunications (TC) verify these guidelines in an effort to find a legal remedy against accused offenders. The results are demonstrated and proved the efficiency of the proposed system by comparing it with the conventional techniques.



***Corresponding Author:***

Imran Ullah Khan
Department of Electronics Engineering,
Integral University
Lucknow, Uttar Pradesh
India
Email: iukhan@iul.ac.in

## 1. INTRODUCTION

Over the course of the past years, crime methods in the telecommunications business have progressed from initial attempts to circumvent payment for communications to more sophisticated schemes involving various participants, systems, and technology. Numerous researches have so far been conducted to analyze, categorize, and offer solutions for the various kinds of fraud[1]. It is more profitable to engage in traffic pumping in nations with high access costs, the connection service prices required by all mobile service providers whenever a call is terminated to another main server. The regulatory authorities set the access rates according to the telecom providers' capacity, request, and innovation. Inadvertently encouraging small operators to artificially produce calls to its system from a major rival is a strong incentive created by induce significant, whose intended function is to prevent price discrimination[2].

When it comes to providing services to customers, telecoms network administrators formerly placed a premium on voice communication due to the high returns it generated. Whether it's wireless or broadband, online communication is becoming the most desirable amenity by customers and businesses together[3].

Nevertheless, telephone is still important and included in many nations' telecommunications packages. Mobile telephone services for consumers and landline telephones for businesses are still provided by operators. The National Regulatory Authorities (NRAs) proceed to fret concerning mobile network regulations including the precise nature of access charges (also called termination rates) . If a call is placed from inside Operator A's (OA) network but is received within Operator B's (OB) network, then Operator B's (OB) network is used to complete the call. This segment of the call costs the accessibility fee or terminating rate (TR)[4]. While telephony has always been an integral aspect of our society, it is now receiving more attention than ever before. Telecommunications companies have a serious issue on their hands due to customer churn, or the number of people who will stop using their services over the next several years. In addition to slowing down the company's expansion, this problem may also reduce earnings. As a result, several customer behaviour prediction models were developed, but they have failed to produce the future outcomes. This is because many of the possible variation factors impacting customer behavior are still being investigated [5].

The practical concern for microgrid applications employing machine learning for intrusion detection. To apply the concepts to various ICSs, the discussion concentrated on a high degree of abstraction. The problems of implementation were then taken into consideration as we spoke about the risks primarily linked to microgrid applications [6]. Thus, the light on the present study in several areas of application-driven on-chip silicon lasers also expresses our opinion on that state of the art. They approach this problem from two different vantage points: the device level and the system-wide viewpoint. The first technique investigates the many pathways that may be followed when integrating on-chip lasers, beginning with the various material systems and on to the integration strategies that can be selected[7]. They examine current and upcoming developments in submarine technology, with a particular emphasis on all of the most important areas, including cable systems, amplifiers' technology, undersea network topologies, electrical power-feeding concerns, economics, and security. The MCF systems are victorious with relative ease over the SC systems. It was also said that submarine-deployed MCFs could have more stringent requirements for the amount of electrical power needed (in comparison to the SMFs) to feed the MC-EDFAs, and they might also have less available physical space[8]. Continuous-wave laser-pumped aluminum gallium arsenide on insulator (AlGaAsOI) nano waveguides make up the spectrum translators. These waveguides provide an octave-wide continuous conversion bandwidth. It is often exceedingly difficult to construct a whole coherent optical communication system for wavelength bands other than C and L due to a shortage of telecom-grade, commercially accessible narrow-linewidth lasers, I/Q modulators, and coherent receivers[9]. An optical signal may be amplified using a remotely pumped erbium-doped system for long-reach (LR) next-generation passive optical networks (NG-PONs). A similar scenario may play out in passive optical networks (PON), which have the largest constraint due to the loss of optical splitters and distribution networks[10]. A convolutional neural network, often known as a CNN, has been developed to identify earthquakes using data that has been collected over three years by fiber cables in communications conduits. The coupling with the ground varies at various segments of the conduits, and human causes such as vehicular traffic create high levels of amplitude noise[11].

## 2. METHOD

As the telecom industry expands at a fast pace, the types of customers that use the service are beginning to vary and diversify. Telecom customer clustering is an efficient strategy for advertising to and keeping loyal customers. Maintaining market dominance requires innovative strategies, and telecom companies have found that segmenting their client base and focusing on retaining certain subsets of consumers is a winning formula. Therefore, it is in the best interest of telecom carriers to investigate effective client clustering algorithms. In this section, we defined the strategic plan. Figure 1 depicts the summary of the proposed method.

### 2.1. Data collection

The research relies on the Cell2Cell dataset, which includes 71,047 instances and 58 characteristics. The information may be found on Kaggle and on the webpage for Duke University's "Centre for Customer Relationship Management." IBM Telco is the second data set for this study. We were able to determine a logical connection between the consumer responses and the turnover by analyzing the information, which contains various variables about a client of a telecoms firm. This dataset can be found on Kaggle and has 21 characteristics and 7044 rows. Several studies have made use of the IBM Telco data set to make predictions about customer attrition. The dependent attribute, "customer churn," is included in both datasets, and may be read as "Yes" or "No," depending on whether or not the consumer has left the organization. Moreover, the characteristics are separated into two types of categories within an object type and one kind of continuous features inside a numerical type in both datasets (Fujo et al. (2022)). Cell2Cell and IBM Telco dataset characteristics are shown in Table 1.

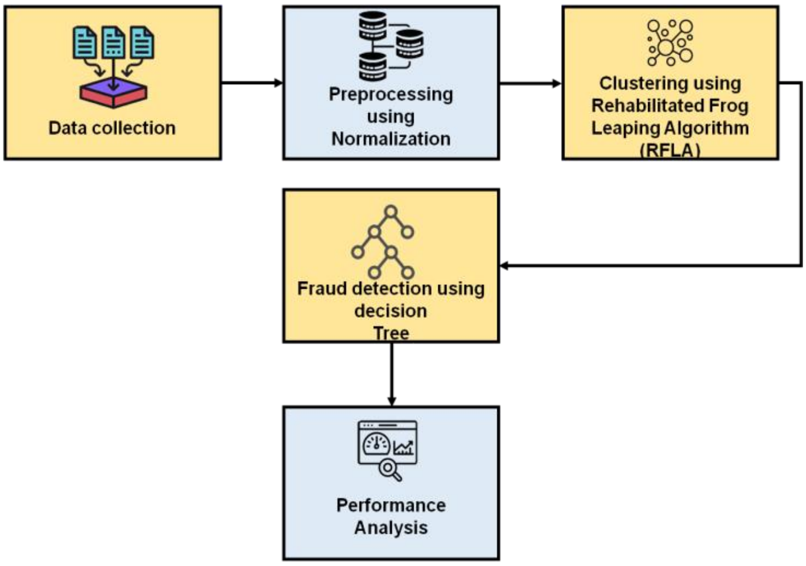


Figure 1. Summary of the proposed method

Table 1. Dataset features

| | IBM Telco | Cell2Cell |
|---|---|---|
| Overall characteristics | 21 | 58 |
| Overall consumers | 7043 | 51048 |
| Error value | Yes | Yes |
| Pumping | 26.5% | 28.8% |
| Not pumping | 73.5% | 71.2% |
| Data dispatched | Imbalanced | Imbalanced |
| Group characteristics | 17 | 23 |
| Statistical characteristics | 4 | 35 |
| Relevant characteristics | 1 | 1 |
| Non-relevant characteristics | 20 | 57 |

### 2.2. Data preprocessing

Data preprocessing refers to every operation done on raw data with the goal to get it ready for further processing. Data normalization refers to the method through which a database's information is made more coherent. This involves constructing records and building links among them according to principles meant to secure the data and increase the database's adaptability by removing unnecessary dependencies and duplications. The information acquired could include important details and irrelevant tidbits, such data packet that have been collected several times. Data cleaning and preparation always includes the elimination of redundant or unused data. It is common practice for academics at analysis organizations to employ estimation techniques because of the large amount of available data. Due to the abundance of distinctive characteristics in this set of data, an extraction of characteristics strategy will be necessary for viewers to filter out irrelevant details. One technique used to ensure uniformity in the data is normalization. This means that it's necessary to standardize the data used for training. The calculation of the z-score is the first step in normalization. To determine a z-score, use Eq. 1. In this equation, where $\sigma$ is the standard deviation, $\rho$ is the average value.

$$Z - score = \left[\frac{\sigma}{(SM - \rho)}\right] \quad (1)$$

$$ZS = \frac{M - \overline{SM}}{sd} \quad (2)$$

After entering into the test statistic $\overline{SM}$ into equation (2), random sample order was described by equation 3, where $\epsilon_a$ denotes sampling error and $SSD^2$ denotes the reliability of the data source.

$$Z_x = \delta_0 + \delta_1 SM_a + \epsilon_a \quad (3)$$

The errors cannot be dependent on each other, as proved by the following equation (4), where $u$ is a free constant.

$$SM_a \sim \sqrt{D}\frac{u}{\sqrt{u^2+D-1}} \tag{4}$$

### 2.3. Clustering using rehabilitated frog leaping algorithm

With little regard for the final result, clustering is a technique for finding and categorizing comparable pieces of data in bigger databases. Data is often categorized via clustering into groupings that are simpler to comprehend and handle. The clustering algorithm creates groups by repeatedly joining the nearest data elements. Each data item is viewed as a distinct group at first since they are all originally separated from one another. The two nearest data elements are then linked together to create a cluster. To develop a way to a sequential optimization issue, the RFLA was developed as a conceptual that employs a heuristic function during an updated systematic search. It is predicated on the spread of ideas between interacting people and a worldwide interchange of material among them.

Assume a bunch of frogs jumping in a swamp. The swamp includes many rocks at different places, and the frogs may jump on any of rocks to pick the one with the most nutrition. The frogs are permitted to converse with one another so they may learn from one another and use that knowledge to better their ideas. When a meme is improved, each frog's location is adjusted to be as advantageous as possible by adjusting how it leaps. By similarity with frogs and its distinct locations, the modification of memotype(s) is still permitted to be a distinct result in this case.

The RFLA combines deterministic and probabilistic methods. The method can effectively employ surface response data to direct the predictive analysis due to the generative model. The searching pattern's adaptability and durability are ensured by the randomized components. A group of frogs chosen randomly covers the whole swamp as the hunt gets under way. The population is divided into a number of separate groups that seem to be free to develop separately and explore the universe in various ways. The frogs undergo computational growth as a result of being acquired by other frogs' views inside each conceptual model. Computational advancement raises the standard of a person's meme and raises that person's effectiveness in achieving an objective. It is necessary that frogs having superior ideas participate more to the creation of ideas than frogs with inferior concepts in order to guarantee that the infected cycle is dynamic. The best knowledge from the whole group may be used by the frogs to alter the memes as they evolve. The new phrase correlates to the frog's new location, and substantial improvements in memotype(s) correlate to a jumping step size. A single frog is reintroduced to the group while making progress in its place. A variation in position results in data that may be used right away to make improvements. This strategy, which needs the whole group to be transformed until unique ideas are offered, differs within the new data is instantly accessible. Once more, the comparison of frogs is utilized, but in regards of the dissemination of concepts, the analogue may be groups of academics or innovators who are developing a notion or designers who are sequentially developing a system.

The promotional are pushed to mingle after a specific period of metaheuristic evolutionary frequently affected, and additional memeplexes are created by a sequence of churning. The memes' performance is improved by this scramble after reinfection by frogs from various parts of the swamp. Frog mobility quickens the searching by spreading their knowledge via infections and ensuring that there is no localized objective in the societal evolution of any specific interest.

#### 2.3.1 Stages of the rehabilitated frog-leaping algorithm

The stages listed below provide a summary of the RFLA approach.

**Global analysis**

Initialize: Choose $m$ and $n$, where $n$ represents the amount of frogs included inside each memeplex and $m$ is the total quantity of memeplexes. As a result, $F = mn$ gives the entire sample size in the swamp.

Create a fictitious group: The possible field $\Omega \subset \mathfrak{R}^d$ contains specimen F interactive frogs $U(1), U(2), \dots. U(F)$ where d is the maximum number of choice variables. An array of choice parameter values $U(i) = \left(U_i^1, U_i^2, \dots. U_i^d\right)$ is used to symbolize the $i^{th}$ frog. Calculate every frog $U(i)$ performance characteristics $f(i)$.

Ranking to frogs: Arrange the F frogs in highest to the lowest of evaluation metrics to rank them. Keep them inside an index with the following formula: $X = \{U(i), f(i), i =$

| | |
|---|---|
| | $1 \dots \dots, F\}$, where $i = 1$ denotes the frog only with highest performance rating. In the overall population of frogs (F frogs), note the highest frog's location $P_X$ (where $P_X = U(1)$). |
| Separating frogs: | Frogs should be divided into iterations. Divide the collection $X$ into m memeplexes $Y^1, Y^2, \dots . Y^m$, each with $n$ frogs, that is $Y^1 = [U(j)^k, f(j)^k \vert U(j)^k = U(k + m(j-1)), f(j)^k = f(k + m(j-1)), j = 1, \dots ., n], k = 1, \dots, m$; (5) |
| Generation of memes inside each memeplex: | Develop every memeplex Based on the frog-leaping method shown below $Y^k, k = 1, \dots \dots, m$. |
| Mix together memeplexes: | Substitute $Y^1, \dots, Y^m$with $X$ so that $X = \{Y^k, k = 1, \dots ., m\}$ after a certain degree of memetic stages of evolution inside each memeplex. Sort $X$ into descending order by evaluation metrics. Modify $P_X$'s optimal frog location for the population. |
| Verify convergence: | If the convergence requirements are met, stop. If not, go back to step 3. Whenever at least one frog continues to carry the "highest metaheuristic sequence" without alteration after a certain number of successive time cycles, the decision to leave is often taken. A limit overall number of computation power may also be established. |

**Local analysis**

Step 0: Set $im = 0$ so that it compares the amount of iterations $im$ represents to the overall number of memeplexes $m$. Set $iN = 0$ so that it compares to the maximal $N$ steps that may be accomplished inside a memeplex. $iN$ measures the amount of evolution process.

Step 1: $Set\ im = im + 1$

Step 2: $Set\ iN = iN + 1$

Step 3: Create a submemeplex:
The frogs want to improve its ideas to grow closer to the best ideas. As previously mentioned, they may borrow concepts from either the top frog in the memeplex $Y^{im}$ or the world's finest. It is sometimes advisable to utilize the best frog when choosing the memeplex since frogs have a propensity to congregate around it specific frog, which could be a localized optima. Consequently, a submemeplex—a portion of the memeplex—is taken into account. According to the submemeplex technique, frogs with significantly better ratings are given more weight, while those who have reduced performance rates are given fewer mass. The mass are distributed according to a triangular probability density function, i.e., $P_j = 2(n + 1 - j)/n(n + 1), j = 1, .. n,$ so that within a group, the frog with the greatest result has the maximum probability to be selected for the submemeplex, with $P_1 = 2/(n + 1)$, and the frog in which the worst effectiveness has the minimum consequence, with $P_n = 2/n(n + 1)$. To create the submemeplex array Z in this case, q different frogs are chosen randomly from the n frogs within every group. Frogs are grouped in the submemeplex in highest to the lowest of efficiency ($iq = 1, \dots q$). In the submemeplex, note the best ($iq = 1$) and worst ($iq = q$) frog positions as $P_B$ and $P_W$ respectively.

Step 4: The worst frog should be placed better: When a frog performs poorly in a submemeplex, the level and specific role are calculated for that frog by

$$step\ size S = min\{int[rand(P_B - P_W)], S_{max}\} \text{ For positive}$$
$$= max\{int[rand(P_B - P_W)], -S_{max}\} \text{ For negative}$$

Here $S_{max}$ seems to be the largest stride size that an affected frog is permitted to assume and $rand$ is a random value between [0, 1]. The step size has proportions that are equivalent to the amount of evaluation criteria, as can be seen. Next, the new location is calculated using

$$U_{(q)} = P_w + S$$

Calculate the new technical value $f$ if $U_{(q)}$ is inside the attainable range $(q)$. If not, go to stage 5. Replace the old $U_{(q)}$ with the fresh $U_{(q)}$ and go to stage 7 if the novel $f_{(q)}$ is superior to the old $f_{(q)}$, i.e., if the evolutionary results in a profit. If not, go to step 5.

Step 5: The frog's new location are calculated for the frog if step 4 is unable to provide an improved outcome.

$$step\ sizeS = min\{int[rand(P_X - P_W)], S_{max}\} \text{ For positive}$$
$$= max\{int[rand(P_X - P_W)], -S_{max}\} \text{ For negative}$$

The above equation is used to calculate the new location. Calculate the new result $f_{(q)}$ if $U_{(q)}$ is in the possible region; else, go to step 6. Replace the old $U_{(q)}$with the fresh $U_{(q)}$ and go to step 7 if the new $f_{(q)}$ is superior to the old $f_{(q)}$, i.e. if the evolutionary results in a profit. If not, go to step 6.

Step 6: Censorship. A fresh frog r is generated randomly at a viable position to replaced the frog whose starting role was unfavorable to advancement if the starting role is either impossible or not better than the previous one. This stops the transmission of the flawed meme. Set $U_{(q)} = r\ and\ f_{(q)} = f(r)$ and estimate $f(r)$.

Step 7: Modernize the memeplex $Z$ should be substituted in $Y^{im}$ for $Z$ once the submemeplex's worst frog has undergone a metaheuristic modification. Sort $Y^{im}$ in descending order for evaluation metrics.

Step 8: If $iN < N$, go to step 2.

Step 9: If $im < m$, go to step 1.

### 2.4. Decision tree

Decision trees are useful for processing sets of data that do not fit neatly into a continuous path. Numerous fields, including architecture, urban planning, legal, and finance, make actual utilisation of decision trees. There are two distinct kinds of decision trees: those based on continuous variables and those based on continuous ones. A tree has several similarities in everyday life, and it turns out to have affected both regression and classification in a broad range of learning algorithms. A decision tree may be used in decision analysis to formally and graphically reflect choices and decision-making.

The decision tree plays a crucial role in translating the grouping model's results into a usable set of guidelines. Because the objective is to expose fraudsters to justice, applicability is not only an important administrative consideration in this situation; it is also a legal limitation. To initiate a case against those responsible for traffic pumping, you must be able to provide convincing evidence of the suspicions. The tree width, pruned intensity, and minimal data are the three parameters that make up a decision tree prediction system. Additionally, a total of 612 decision rules are developed.

## 3. RESULTS AND DISCUSSION

We compare the rehabilitation frog leaping algorithm (RFLA) to modern techniques like Big Data, Broad Reinforcement Learning (BRL), and Deep Learning (DL) to categorize potential fraud instances into separate categories. The findings demonstrate that the suggested technique outperforms the state-of-the-art in terms of Call length, Call category, Disposition, Error, and Business Hour.

### 3.1. Call Duration

Figure 2 shows the length of the call. It is the duration of a phone conversation between two persons or the period during which information is transferred. Direct listening to digital or analog voice communication, as well as data about any kind of communication being intercepted or sniffed.

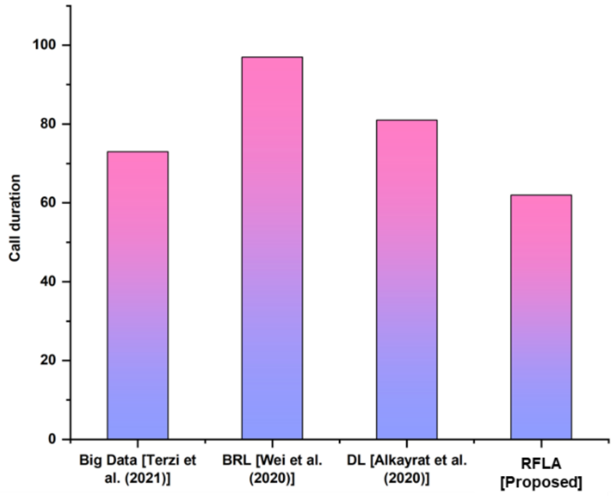


Figure 2. Call Duration

The call duration for the current approaches, which were big data (73%), BRL (97%), and DL (81%), and the suggested system, was 62%. This suggests that the proposed system is more efficient. The call duration is shown in Table 2.

Table 2. Call Duration

| | Call duration |
|---|---|
| Big Data [Terzi et al. (2021)] | 73 |
| BRL [Wei et al. (2020)] | 97 |
| DL [Alkayrat et al. (2020)] | 81 |
| RFCA [Proposed] | 62 |

### 3.2. Disposition

Figure 3 illustrates a call disposition and provides information on the result of a call. Statuses such as "demo planned," "left a voicemail," and even "no longer in service" are among them. In point of fact, customer support teams have been logging call dispositions for some years, and sales teams have now followed suit and begun recording call outcomes.

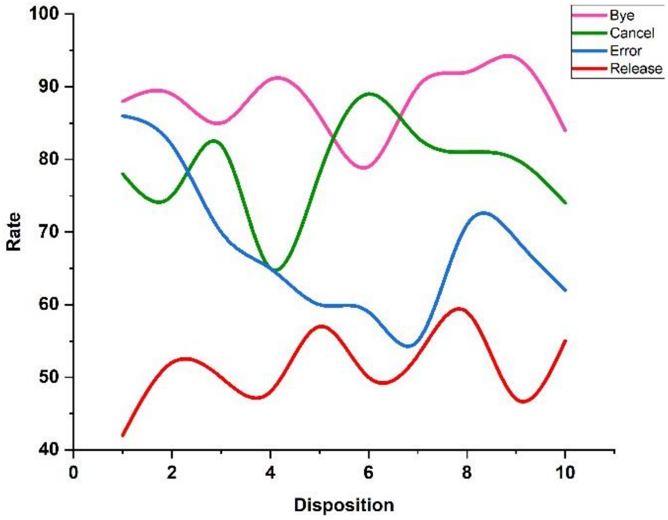


Figure 3. Disposition

To ensure that all members of your team are on the same page, you may classify both incoming and outgoing calls by using call disposition codes. The present standing or ultimate conclusion of an arrest or prosecution is referred to as the disposition of a criminal record.

### 3.3. Business Hour

A business hour is shown in Figure 4; during normal business hours, you may expect to have your phone answered by a real person. Create a timetable that is simple to follow. The hour of the month in which a certain telecommunications system transports the most traffic. The 60-minute period during which the greatest overall traffic load in a specific 24-hour period occurs.

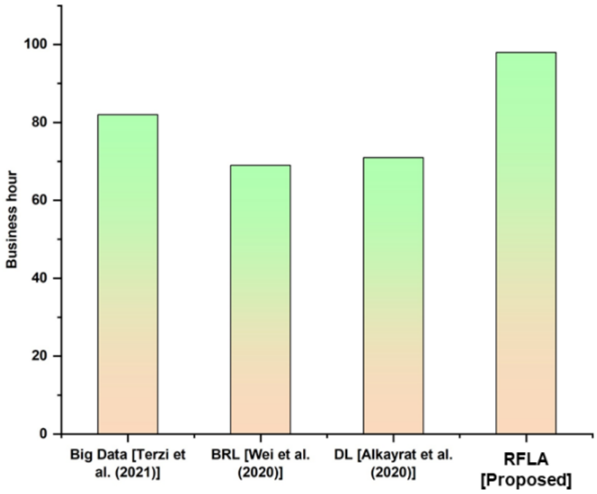


Figure 4. Business Hour

The business hour for the present techniques were big data (82%), BRL (69%), and DL (71%), as well as the recommended system, which was 98%. This lends credence to the idea that the suggested approach is more effective. The business hour is shown in Table 3.

Table 3. Business Hour

| | Business hour |
|---|---|
| Big Data [Terzi et al. (2021)] | 82 |
| BRL [Wei et al. (2020)] | 69 |
| DL [Alkayrat et al. (2020)] | 71 |
| RFCA [Proposed] | 98 |

### 3.4. Day Frame

Computer networking and communications make use of something called a day frame, which is a digital data transmission unit. This notion is shown graphically in Figure 5. A simple container for a single network packet is what's referred to as a "frame" in computer networks that makes use of packet switching. A frame is a time-division multiplexing (TDM) supporting structure that may be found in different types of telecommunications systems.

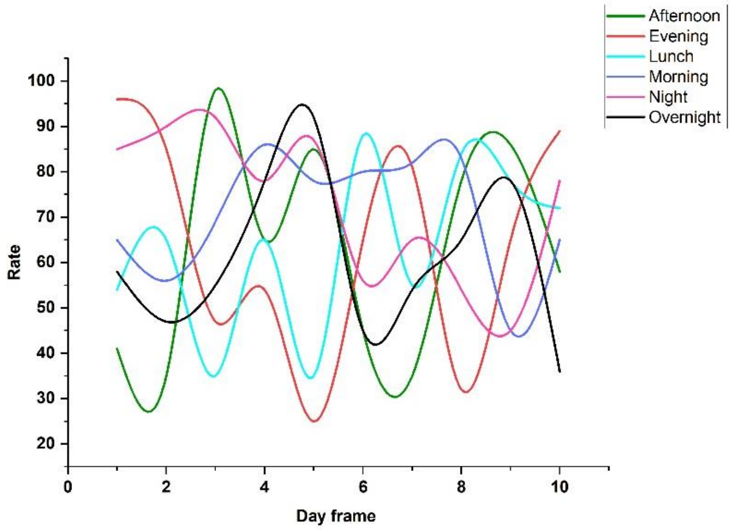


Figure 5. Day Frame

### 3.5. Error

Figure 6 is a representation of the error, which is a measure of the degree to which a model's predictions deviate from those of the actual model. The error rate is determined by taking the total number of bits that were transferred in a certain amount of time and dividing that number by the number of bits that were received erroneously. When it comes to telecommunications, an error rate of 10-9 is typically believed to be acceptable, however, a minimum error rate of 10-13 is thought to be more suited for data transfer.

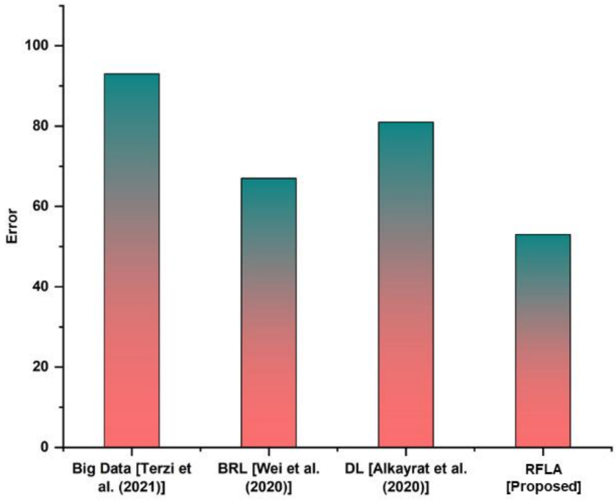


Figure 6. Error

The error for the present techniques was big data (93%), BRL (67%), and DL (81%), as well as the recommended system, which was 53%. This lends credence to the idea that the suggested approach is more effective. The Error is shown in Table 4.

Table 4. Error

| | **Error** |
|---|---|
| Big Data [Terzi et al. (2021)] | 93 |
| BRL [Wei et al. (2020)] | 67 |
| DL [Alkayrat et al. (2020)] | 81 |
| RFCA [Proposed] | 53 |

**Discussion:**

Big data analytics, in comparison to more conventional methods, generate proactive models, freeing them from the constraints of analysis and storage. These models also provide insights that can be acted upon, thereby assisting in the improvement of business processes and the quality of relationships with customers. Many challenges had to be overcome to develop the model and obtain the results as a real-world application, including limitations on non-disclosure agreements, dynamic fraud patterns, restrictions on data intervals and frequency, challenges in obtaining labeled data, irregular distributions of both fraud and legitimate classes, challenges in analyzing the data in distributed file systems (Terzi et al. (2021)). To get the most out of an environment, one must learn the best conduct to exhibit. Decision-making science is what it is. To get the most out of an environment, one must learn the best conduct to exhibit. Overload and a weaker outcome might arise from too much reinforcement. When an activity is aborted or avoided as a result of a bad circumstance, that reinforcement is deemed negative (Wei et al. (2020)). Deep learning-based algorithms have shown significant results in the removal of duplicated and unnecessary data. The objective of data representation and clustering is often achieved by academics using autoencoders and clustering techniques, which is made possible by the widespread interest in deep learning. Instead of supporting merely linear data transformation like other fundamental methods like PCA, autoencoders may learn data projections with the right dimensions and sparsity restrictions (Alkhayrat et al. (2020)). This method permits message exchange in local search and employs a hybrid technique. The benefits of the Nomometric method and particle group optimization are combined in this algorithm. When compared to other existing technologies, RFLA is more effective.

## 4. Conclusion

Local telephone operators increase the quantity of incoming calls to own systems in a practice known as "traffic pumping" in order to benefit from a relevant access price than the wireless carrier linked to the call's origin. This study describes a clustering-based decision support system for dealing with traffic pumping. Despite the fact that, numerous research highlights the significance of this scam kind. To the authors' knowledge, our suggestion is predicated on the clustering strategy designed for traffic pumping. By examining the findings, the suggested system's effectiveness is shown. In relation to the telecommunications scam including traffic congestion, the inaccuracy might be estimated at 53%. In comparison to other features, the suggested method has a 98% efficiency rate during business hours when measuring fraud activities. The ideal learning strategies may be used to illustrate the research's potential future use.

**REFERENCES**


[[1] Chadwick, S.R., Grinstein, S. and Freeman, S.A., 2021. From the inside out: Ion fluxes at the centre of endocytic traffic. Current Opinion in Cell Biology, 71, pp.77-86.

[2] Galinmoghadam, J. and Zhang, X., 2022. Performance of Wicking Geotextile (H2Ri) To Mitigate Pavement Pumping–Phase 2 (No. cmr 22-008). Missouri. Department of Transportation. Construction and Materials Division.

[3] Irarrázaval, M.E., Maldonado, S., Pérez, J. and Vairetti, C., 2021. Telecom traffic pumping analytics via explainable data science. Decision Support Systems, 150, p.113559.

[4] Lacuška, M. and Peráček, T., 2021. Trends in global telecommunication fraud and its impact on business. Developments in Information & Knowledge Management for Business Applications: Volume 1, pp.459-485.

[5] Chadyšas, V., Bugajev, A., Kriauzienė, R. and Vasilecas, O., 2022, June. Outlier analysis for telecom fraud detection. In Digital Business and Intelligent Systems: 15th International Baltic Conference, Baltic DB&IS 2022, Riga, Latvia, July 4–6, 2022, Proceedings (pp. 219-231). Cham: Springer International Publishing.

[6] Ramotsoela, D.T., Hancke, G.P. and Abu-Mahfouz, A.M., 2023. Practical Challenges of Attack Detection in Microgrids Using Machine Learning. *Journal of Sensor and Actuator Networks*, *12*(1), p.7.

[7] Bose, S. Subash Chandra, Badria Sulaiman Alfurhood, Gururaj H L, Francesco Flammini, Rajesh Natarajan, and Sheela Shankarappa Jaya. 2023. "Decision Fault Tree Learning and Differential Lyapunov Optimal Control for Path Tracking" Entropy 25, no. 3: 443. https://doi.org/10.3390/e25030443

[8] Papapavlou, C., Paximadis, K., Uzunidis, D. and Tomkos, I., 2022, June. Toward SDM-based submarine optical networks: A review of their evolution and upcoming trends. In Telecom (Vol. 3, No. 2, pp. 234-280). Multidisciplinary Digital Publishing Institute.

[9] Kong, D., Liu, Y., Ren, Z., Jung, Y., Kim, C., Chen, Y., Wheeler, N.V., Petrovich, M.N., Pu, M., Yvind, K. and Galili, M., 2022. Super-broadband on-chip continuous spectral translation unlocking coherent optical communications beyond conventional telecom bands. Nature Communications, 13(1), p.4139.

[10] Munster, P., Horvath, T., Dejdar, P. and Skaljo, E., 2022. Remote Pumping in Point to Multi-point Networks. Journal of Communications Software and Systems, 18(1), pp.54-62.

[11] Huot, F., Biondi, B.L. and Clapp, R.G., 2022. Detecting local earthquakes via fiber-optic cables in telecommunication conduits under Stanford University campus using deep learning. arXiv preprint arXiv:2203.05932.

[12] Fujo, S.W., Subramanian, S. and Khder, M.A., 2022. Customer churn prediction in telecommunication industry using deep learning. Information Sciences Letters, 11(1), p.24.

[13] Terzi, D.S., Sağıroğlu, Ş. and Kılınç, H., 2021. Telecom fraud detection with big data analytics. International Journal of Data Science, 6(3), pp.191-204.

[14] Wei, X., Zhao, J., Zhou, L., and Qian, Y., 2020. Broad reinforcement learning for supporting fast autonomous IoT. IEEE Internet of Things Journal, 7(8), pp.7010-7020.

[15] Alkhayrat, M., Aljnidi, M. and Aljoumaa, K., 2020. A comparative dimensionality reduction study in telecom customer segmentation using deep learning and PCA. Journal of Big Data, 7, pp.1-23.

**BIOGRAPHIES OF AUTHORS**

**Imran Ullah Khan** is Associate Professor in the Department of Electronics Engineering Integral University, Lucknow. He has more than Twenty Years teaching experience. He has completed B. Tech and M. Tech from Aligarh Muslim University (AMU), Aligarh and Ph.D. from Mewar University, Chittorgarh. Ph.D. in Electronics and Communication Engineering and I have 20+ years of relevant Academic Experience. I am University Center Coordinator of Indian Institute of Remote Sensing, Indian Space Research Organization, department of space, government of INDIA from October, 2015. Executive Committee Member (Ex) IEEE UP Section. Ambassador IEEE DAY-2017 INDIA, Pakistan and Bangladesh. Worked as Head Electronics & Communication Engineering Department in Maharana Institute of Professional Studies, Kanpur from 10-01-2012 to 21-07-2013.Worked as Controller of Examination in Maharana Institute of Professional Studies, Kanpur from 10-01-2012 to 21-07-2013.Worked as Dean Students Welfare in Maharana Institute of Professional Studies, Kanpur from 10-01-2012 to 21-07-2013. Worked as Head Electronics & Communication Engineering Department in Kanpur institute of Technology, Kanpur from 02-07-2010 to till 09-01-2012. Outstanding Branch Counselor Award during Annual General Meeting held on January 29, 2017 at IIT, Kanpur for the exemplary services as Integral University Branch Counselor of IEEE UP Section for year 2016. Conference Chair and Organizing Secretary "First IEEE International Conference on Computational and Characterization Techniques in Engineering & Sciences" (CCTES-18) (IEEE Conference #44023) scheduled on September 14-15, 2018. Worked as Organizing Secretary in International Symposium on "Computational and Characterization Techniques in Engineering & Sciences" (CCTES-17) held on March 20, 2017. Worked as Organizing Secretary in International Seminar on "Present Scenario & Future Prospectives of Research in Engineering and Sciences" (ISPSFPRES-17) held on January 21, 2017. Reviewed "Electronic Communication Systems" by George Kennedy and Bernard Davis with Tata McGraw Hill Education Private Limited, New Delhi. He can be contacted at email: iukhan@iul.ac.in

**Naga Lakshmi Sowjanya Cherukupalli** is Assistant Professor, CSE Department, Koneru Lakshmaiah Education Foundation, Green Fields, Vaddeswaram, Guntur, India and She is also a Research Scholar in Gitam Institute of Sciences, Vishakhapatnam, India. She has published many national and international conference papers and as well as many scopus and sci journal papers. She has published many Indian innovative patents in the field of computer science and engineering. She has also attended many workshops, seminars, and training in the same field. She can be contacted at email: chlks1709@kluniversity.in