

# Self-Organizing Map based Wallboards to Interpret Sudden Call Hikes in Contact Centers

Samaranayaka J. R. A. C. P<sup>1</sup>, Prasad Wimalaratne<sup>2</sup>  
University of Colombo School of Computing  
Colombo, Sri Lanka

**Abstract**—In a contact center, it is required to foresee and excavate any disturbance to the daily experiencing call pattern. Abnormal call pattern may be a result of a sudden change in the organization's external world. Expecting a methodological analysis prior to meet customers' demand may introduce a delay for queuing customers. It is required a fast and promising method to predict and reasoning any unwilling event. It is not possible to draw conclusions by considering one dimension such as total call count. Total call count may increase in same way due to a failure in any service. Research mainly focusses on reasoning multidimensional events based on historical records. In contrast to traditional wallboards, our approach is capable of clustering and predicting disturbances to the normal call patterns based on historical knowledge by considering many dimensions such as queue statistics of many service queues. Our approach showed improved results over traditional wallboards equipped with 2D or 3D graphs.

**Keywords**—Multidimensional data; visualization; contact centers; self-organizing map; clustering

## I. INTRODUCTION

Contact center is a very sensitive customer touch point within an organization [1]. In a different explanation, any sudden incident in the external world of an organization will be reflected immediately via contact centers. Contact center is a combination of both technology and human who are waiting to provide information required by the callers. Neither customer care officer nor a customer will be able to conclude a sudden change in the outside world of an organization. Simply, incident can be exemplified as an electricity failure of a region for national electricity provider or common telephone cable damage due to a heavy rain for national telephone provider. In both cases, customers may unhappy about the service interruption and they will try to complain about it by resulting a huge demand upon the relevant contact centers.

It will create a need for a software mediator which could extract information from both caller and customer care office, analyze against historical knowledge and present to the supervisors who can decide any deviation from normal operational procedures. Such an operational deviation is mandatory in a sudden event because contact centers were not designed to absorb abnormal conditions to associated capital investments. Financially benefited strategy is maintaining a separate approach for sudden incidents. To have a competitive advantage over the competitors who are in the same field, it is required a good insight about when to switch strategies without diluting customer satisfaction. Our research focusses on introducing wallboards with capability to

visualize multidimensional data along with clustering to support real time decision making with the support of historical knowledge.

## II. RELATED WORK

Due to increasingly generating and accumulating large amount of unstructured data, it is often tending to excavating knowledge behind the data. While analyzers interested about user-friendly interactive systems for knowledge extraction, managers like people who are working in a contact center or help desk are more interested about wallboard [2] type information systems to alert timely any deviation with related to their predictions or organizational directions. This para will motivate the trending research interest about visualizing dynamic data. Galkin and others have examined a pipeline approach to visualize and analyze dynamic data to facilitate anomaly detection, clustering, trend analysis and variation analysis [3]. Their research explained that analyzing complex multidimensional data will introduce a deep insight to data science and related industries. This can be accomplished by using the influence of trending technologies such as artificial intelligence, machine learning and neural networks [4], [5]. In the past, several research interests have been showed in the area of time dependent multidimensional data visualization [6], [7] and one of the common observation in the above mentioned area is seeking the assistance of a domain expert to interpret the result which was projected to a 2D or 3D coordinate system. In Steven's approach, they are using a progressive technique which rendering the incoming data progressively. In other words, framework will initially construct a geometry to represent a past summery and updating a dynamic scalar field on top of that geometry [8]. In their research, they have tried to address a challenge which is handling and representing dynamic scalar in an efficient and user friendly manner. Mashima and other have presented a map based metaphor for visualizing dynamic data [9], [10]. Point (peg) in a map represents a multidimensional data while distances from the other points on the map are proportional to the similarity between points. In the other words, higher the distance mean less similar points. Researchers suggested that animations can be used as user interactions to successfully present dynamic scalars in a view. Wallboard kind of applications can be equipped with periodic updates or animations to successfully project multidimensional data with dynamic behavior to a 2D screen.

Throughout the history data visualization was a trending topic due to need of knowing "what is it hiding". People tried to interpret hidden knowledge in a human understandable way

[11]. A better visual representation is easily understandable and holds features like effectiveness, accuracy, robustness, easy to use, etc. [12], [13]. Historical effort for visualizing multidimensional data can be exemplified with techniques like Scatter Plot matrix. With increase of velocity, volume and veracity of data collections, it opened new research areas like virtual reality for data visualization, augmented reality for data visualization, effective use of interaction methods for data visualization, etc. [14]- [16]. Wallboard is a display placed in a public area to convey any interested information. It can be equipped with effective interaction techniques.

Another important branch of visualization is dimensionality reduction techniques such as Principal Component Analysis (PCA) [17], Radial Coordinate visualization (RadViz) [18], self-organizing maps (SOM) [19], Ridge and Lasso Regression [20], [21], Singular Value Decomposition (SVD) [22]. Above two paragraphs and techniques were motivated to build a hybrid visualization model which is a collection of both multidimensional visualization models' characteristics and dimensionality reduction techniques' characteristics. Among the dimensionality reduction techniques, self-organizing maps showed better performance for clustering over other clustering techniques [23].

Interactivity can be embedded to a concept by using technologies like JavaScript, Ajax, HTML, Hadoop [24], NodeJS, etc. These technologies collectively introduce continuous connectivity in-between frontend and backend for enabling real time or near real time updates in the front end.

Since visualization is for humans, it is needed to consider human factors in visualization [25], [26]. Ability to operate in a low resources environment is a greater achievement with increasingly generating various types of data. This fact led research to design and implement a distributed architecture for information processing.

### III. WALLBOARD

Below will demonstrate commonly available wallboard types with both open source and commercial contact center products. Asterisk can be referred under the free and open source contact centers. They may not provide identical wallboards as below figures (Fig. 1 and Fig. 2). But, it will maintain the same concept.

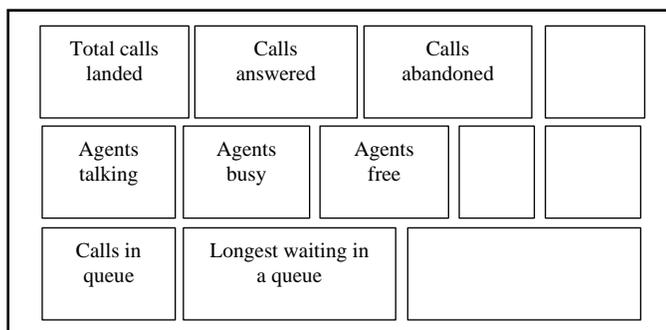


Fig. 1. Standard Wallboard of a Contact Center.

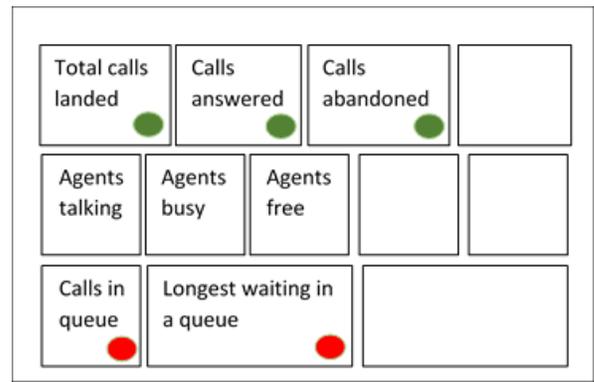


Fig. 2. Standard Wallboard of a Contact Center with Alarms based on Past Data (it will Indicate Whether a shown Value Exceeds a Threshold or Not).

Because of market competition, both open source and commercial contact center products tend to offer more informative dashboards or information visualization modules with their products. Presently, pattern recognition is a lacking and demanding area with information visualization which is related to contact center solutions.

### IV. METHOD

#### A. Customer Information flow of a General Contact Center

Generally, contact center will consist with below major logical modules to cater demanding customer requirements.

- Greeting message: This is the first message customer can here. Example: Good morning.
- Skill: Skill is an attribute which is assigned to an agent. Agent Skills can be regard as the ability of an agent to handle a specific call which requires one of those skills. In relationship with contact center, skill can be thought as a specific customer need/requirement or perhaps a business need of contact center. Contact center will define skills based on the needs of customers and contact center.
- Interactive Voice Response (IVR): It provides a list of steps that process calls in a user-defined manner. The steps in an IVR can send calls to skill, play announcements and music, disconnect calls, give a busy signal to the calls, or route calls to other destinations.
- Queue: Queue is a holding area for calls which are waiting to be answered in order. Different calls in a queue may have different priority levels, in which case, calls with a higher priority will be answered first.
- Agents terminal: This is the place were agents can pick and answer the customer's call. Based on the organization's requirement calls may be landed automatically to the relevant agent's terminal.

#### B. Placement of the Analyzer

As above diagram (Fig. 3) shows, information flows through the above explained logical modules. Skill was embedded to agent who is using agent terminal and it is not separately drawn in the diagram. Selection of the information

flow which is needed to process was done based on below facts (Table I).

- Relevance to the presented analysis
- Richness of the information
- Experience of contact center management
- Experience of number management teams (Soft switch, etc.).

Based on the comments of number management team, due to central signal handling of phone calls, it introduces errors when grouping numbers into a geographical region by using number levels. For an example, number level 1111xxxxx – 11119xxxxx is belong to town A and number level 11121xxxxx – 11129xxxxx is belong to town B. Town A and B belongs to region X, since 111xxxxxxx assigned to region X. Because of the central management of signals (central switch), a number belongs to town A can be assigned to region Y. Because of this configuration, it dilutes the capability of localizing a regional failure due to a common reason which occurred in a particular region such as heavy rain.

Based on the experience of contact center management, contact center will be flooded from calls due to a failure in a region. In a such situation, we cannot experience a patient behavior from callers. All of the callers may have to wait longer minutes to reach agents. Most interesting observation is that caller is not capable to predict the source of failure and they will report what they feel. Because of that caller may select the required service based on their perspective. In this research, most significant information to analyze is queue statistic over the other available information when experiencing a sudden call hike. Based on the above explanation, we have selected information flow between queue and agent terminal to analyze.

C. Work Flow of the Proposed Software

Work flow (Fig. 4) was implemented with our solution to visualize results. As described by following figure, it will periodically calculate SOM and visualize its results. On top of the SOM’s visualized content, it will place the present reading’s 3D graphics. This is a periodic process with an appropriate delay to update visualized content or graphics.

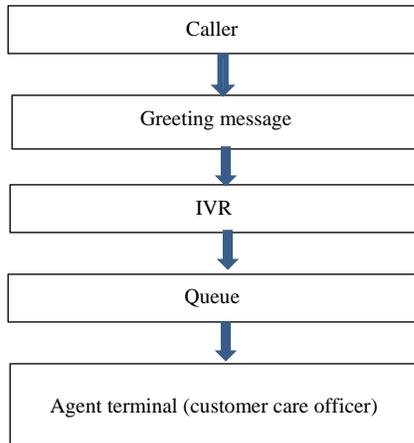


Fig. 3. Customer Information flow of a General Contact Center.

TABLE I. INFORMATION FLOWS BETWEEN COMPONENTS OF A GENERAL CONTACT CENTER

Information flow	Available information (richness)	Relevance to the presented analysis
Caller to greeting message	CLI (caller’s telephone number)	Yes
Greeting message to IVR	CLI (caller’s telephone number)	Yes
IVR to Queue	CLI, Customers required service	Yes
Queue to Agent	CLI, Customers required service, Queue statistics (waiting time, etc.)	Yes

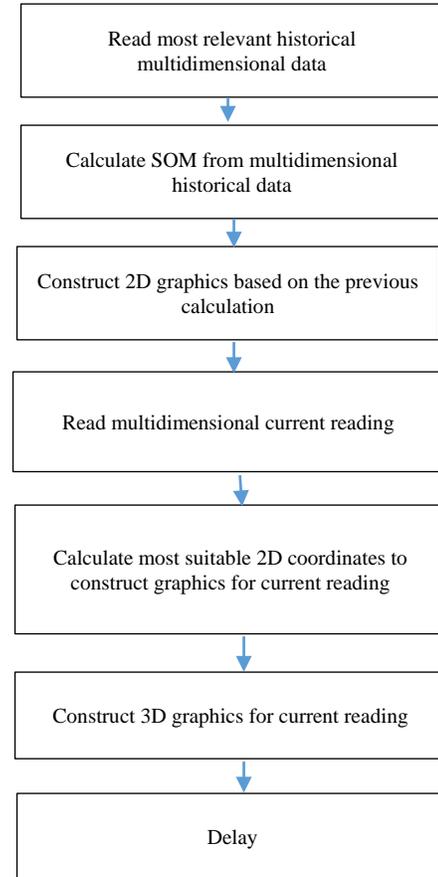


Fig. 4. Proposed Approach’s Dynamic Data updating Work Flow.

D. Architecture of the Proposed Software

Below architecture (Fig. 5) consisted with two main modules such as Visualization layer and “SOM algorithm implementation” module. They are separately described in below subsections in detail. “DB interface” was designed to insert data originated via contact center. Below explained architecture was implemented by using multiple technologies collaboratively such as C#, PHP, MySQL, JavaScript, Three.js, WebGL, Java, etc.

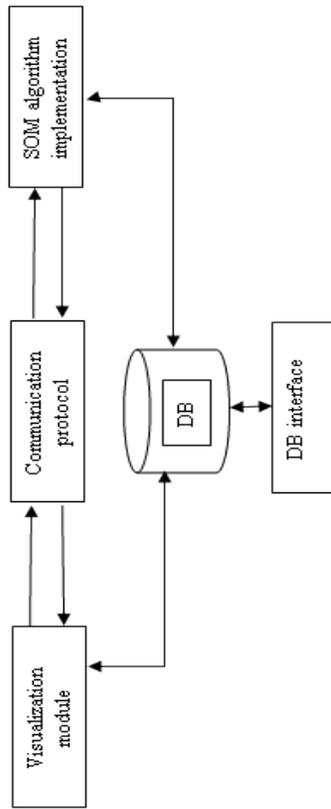


Fig. 5. Architecture of the Proposed Software.

## V. THEORY AND MODULES

### A. “SOM Algorithm Implementation” Module

Kohonen Self-Organizing Map follows a two-layer approach which consists with input layer (top layer of above diagram (Fig. 6)) and output layer (bottom layer of above diagram (Fig. 6)). Both layers’ neurons have same number of dimensions like  $(w_{ij1}, w_{ij2}, w_{ij3}, \dots, w_{ijn})$  and  $(x_1, x_2, x_3, \dots, x_n)$ . Output layer’s neurons’ weights will be updated by using a neighborhood function and input layer’s weights. As per the algorithm, input layer’s neurons will be selected randomly and iteratively to update output layer’s neurons via neighborhood function. Once the input layer’s neuron was selected, matching output layer’s neuron will be selected as per the below function.

$$D_{ij} = |X - W_{ij}| = \sqrt{(x_1 - w_{ij1})^2 + \dots + (x_n - w_{ijn})^2} \quad (1)$$

Here, i and j: output layer indexes (row number and column number)

Closest node,

$$D(k_1, k_2) = \min D_{i,j} \quad (2)$$

$k_1$  and  $k_2$  : winning node’s indexes

Secondly, below equations will explain the neighborhood function which will be used to update neighbors and its calculations.

$$h(\rho, t) = \exp\left(\frac{\rho^2}{2\sigma^2(t)}\right) \quad (3)$$

$$\rho = \sqrt{(k_1 - i)^2 + (k_2 - j)^2} \quad (4)$$

Here i and j: output layer indexes

$k_1$  and  $k_2$  : matching node’s indexes

$$\sigma(t) = \sigma_0 \exp\left(\frac{-t}{T_{max}}\right) \quad (5)$$

t: 0 ..  $T_{max}$

$\sigma_0$ : experimental constant.

### B. Visualization Module

Cartesian coordinate system or spherical coordinate system is capable of handling maximum three variables at a time. We have combined two coordinate systems in our novel framework to enable plotting more than three variables simultaneously. Advantages of this technique over other available 3D visualization approaches are as follows.

- Consistency of all the plotted parameters/dimensions
- Ability to compare and contrast its own dimensions

Above explanation will elaborate in detail in the results and analysis section by using screenshots of the novel framework. In other words, this is the module which visualizes the output of “SOM algorithm implementation” module. This is an independent software and visualization module can be used to upload data by using data files, if it is not integrated with any application to pump data.

### C. Communication between Modules

Communication between visualization layer (web application) and “SOM algorithm implementation” module was facilitated by this layer. As below diagram (Fig. 7) explained, this layer is a composition of TCP socket server and TCP socket client which were constructed as independent executables with separate configuration files. These executables will exchange metadata and results between above mentioned two modules. Transmitting preprocessed data or SOM algorithm’s output is excluded from communication layer protocol and they will be placed on a shared location or can be transmitted by using FTP based on requirement.

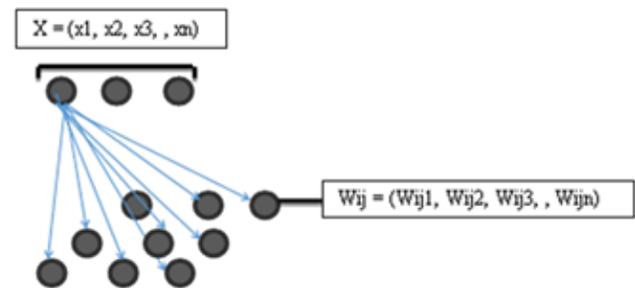


Fig. 6. Diagram of Kohonen Self-Organizing Map.



Fig. 7. Placement of TCP Socket Server and TCP Socket Client in Modules.

## VI. INTEGRATION BETWEEN CONTACT CENTER AND PROPOSED SOFTWARE

As per the below diagram (Fig. 8), contact center's Queue module can be integrated with "DB interface" to pump queue statistics data to database for the purpose of visualization and decision making. Upon successfully reaching an agent collected queue statistics will be sent to database via "DB interface". Then, visualization module can be loaded by using a preferred web browser to view output. It can be used as a wallboard to continuously monitor and analyze call patterns.

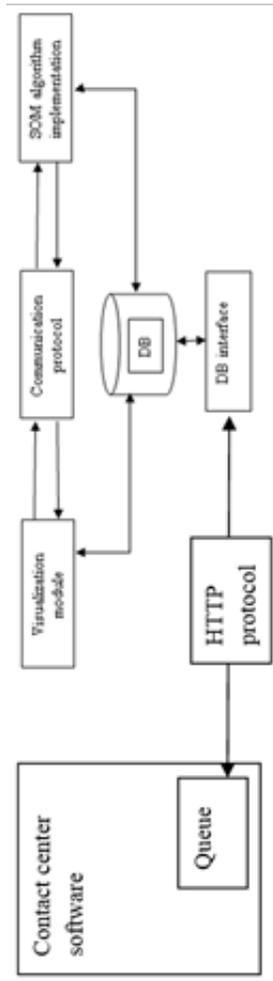


Fig. 8. Diagram of Integration between Contact Center and Proposed Software.

## VII. DATASET AND DIMENSIONS

This section will explain the dataset which can be introduced as queue statistics and its structure in detail. Mainly it has two tables (Table II and Table III). Answered call load pattern table was filled with successfully answered calls by contact center agents while abandoned call load pattern table was filled with calls which is not answered by a contact center agent. Answered call load pattern table has columns like,

- (0-5) which means number of callers who had to wait in a queue between 0 to 5 seconds before answered by a contact center agent in a given month.

- (6-10) which means number of callers who had to wait in a queue between 6 to 10 seconds before answered by a contact center agent in a given month.
- (11-15) which means number of callers who had to wait in a queue between 11 to 15 seconds before answered by a contact center agent in a given month.
- (>60) which means number of callers who had to wait in a queue more than 60 seconds before answered by a contact center agent in a given month.

Other than monthly cumulative call counts, it is possible to calculate cumulative figures based on 15-minute duration, hourly, daily and monthly. Our proposed software can accommodate all of the above mentioned cumulative figures. We have used daily and monthly cumulative figures to verify and calculate our approach's accuracy.

There may be many reasons to experience a sudden call hike in a contact center such as,

- Credit control action. Operator will discontent customers who did not pay the bill as a bulk. In such a situation, customer may try to reach contact center by asking information related to his/her bill.
- Regional service failure. In such a situation, large number of affected customers will try to complain about service interruption via contact center.
- Natural event such as heavy rain. This kind of events may lead to service interruptions, because of damage to outside infrastructure and delayed maintenance.

We have selected to analyze rainy seasons with related to call hikes. This selection was solely made to evaluate our approach's accuracy. But, our approach can be used with any of the above mentioned reasons. Rain is a seasonal event and it was not started and ended in a same day of a month in two different years. Monthly cumulative call counts will provide better accuracy which is more than 95%.

Same dataset can be used with different durations to calculate cumulative figures to detect any interested reason which is contributing to a sudden call hike.

In contrast with answered call load pattern table, abandoned call load pattern table has columns like,

- (0-5) which means number of callers who had to wait in a queue between 0 to 5 seconds before call is dropped in a given month.
- (6-10) which means number of callers who had to wait in a queue between 6 to 10 seconds before call is dropped in a given month.
- (11-15) which means number of callers who had to wait in a queue between 11 to 15 seconds before call is dropped in a given month.
- (>60) which means number of callers who had to wait in a queue more than 60 seconds before call is dropped in a given month.

TABLE II. ANSWERED CALL LOAD PATTERN

Month	0-5	6-10	11-15		51-55	56-60	>60
01/2015	302708	13354	9994		6995	7161	158597
01/2016	405035	9517	5352		3269	3170	44328
01/2017	456929	10853	8063		4384	4131	80207
01/2018	473343	10302	6471		3185	3064	38553

TABLE III. ABANDONED CALL LOAD PATTERN

Month	0-5	6-10	11-15		51-55	56-60	>60
01/2015	6952	3683	2520		2429	9001	88680
01/2016	3793	1652	849		830	3545	18691
01/2017	5507	2344	1666		1956	5716	42317
01/2018	3839	1277	718		582	3643	14399

Same as explained with answered calls, it is possible to calculate cumulative figures based on 15-minute duration, hourly, daily and monthly for abandoned calls. Our proposed software can accommodate all of the above mentioned cumulative figures.

Dimensions of our multidimensional data instances will be columns of both answered call load pattern table and abandoned call load pattern table or columns of abandoned call load pattern table based on the scenario.

- Scenario 1: if the impact of the event is short duration (several minutes) and contact center employed extra agents to tolerate sudden hikes, both answered call load pattern table and abandoned call load pattern table can be considered as the input for our approach.
- Scenario 2: if the impact of the event is long duration (several hours or days) and contact center has no extra agents to tolerate sudden hikes, abandoned call load pattern table can be considered as the input for our approach.

Rain is a seasonal event and it will introduce a long term impact to a contact center. In our analysis, it was confirmed that there are no additional agents to flatten sudden hikes. In such a scenario, there is no considerable variation with the values of answered table while there is a considerable variation with the values of abandon table. This will be explained in detail in results and analysis section. Considering more dimensions for the research will introduce a requirement for high processing power with related to both calculations and graphics. In this research, we have selected abandoned table's columns as dimensions for our multidimensional data instances.

By reducing the duration for previously explained cumulative value calculations, dashboards can be converted to real time sudden call hike indicators. Although our proposed approach is independent from the dataset (our novel approach is generalized and we have used this dataset to verify and calculate our approach's accuracy which is more than 95%), domain knowledge is mandatory to select most appropriate dimensions and durations for an analysis.

### VIII. RESULTS AND ANALYSIS

In this section, we are describing the results drawn via our novel approach which was described along with the dataset throughout several sections. We have selected queue statistics from historical data DB randomly for rainy months and normal months. Randomly selected data (multidimensional data arrays which was explained in previous section) were categorized by using SOM algorithm and visualized as below (Fig. 9) (blue and green clusters for rainy and normal months).

Workflow of the proposed software was explained in a previous section. According to the workflow, next step is placing the current value which is a multidimensional data instance on top of the visualized SOM. Three white spheres imply three tests for the proposed application and it was verified that drawn conclusions against actual records. A sample of tests which were conducted to verify shown results in the visualization layer were presented in below table (Table IV). It was achieved more than 95% accuracy level for detecting rainy conditions which may influence sudden call hikes. Although this testing was conducted to calculate accuracy and verify functionality of the proposed software, this approach was generalized to use with any different dataset and scenario.

White spheres consist with yellow bars which represent magnitudes of its own dimensions on its perimeter. Our concept's visualization layer was innovated to visualize multidimensional data with the help of SOM based categorization. Table V shows a comparison between our novel visualization concept vs groups of existing multidimensional visualization techniques.

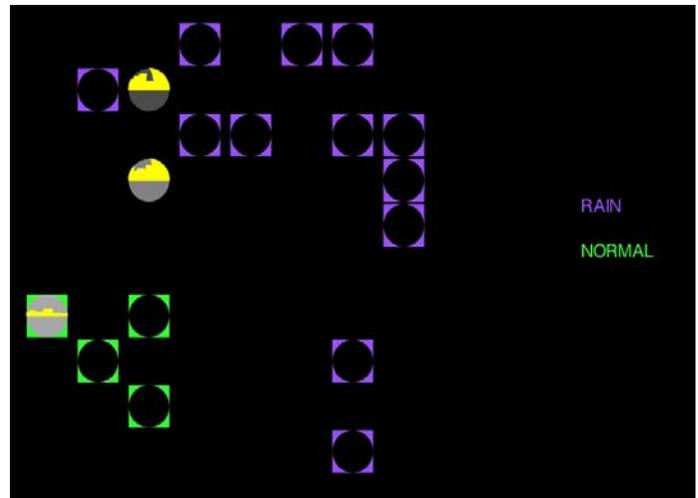


Fig. 9. Output of the Novel Application.

TABLE IV. SAMPLE OF TESTS CONDUCTED WITH VISUALIZATION LAYER

Number	Rainy or normal according to visualized output	Rainy or normal according to past records
1	Normal	Normal
2	Rainy	Rainy
3	Rainy	Rainy

TABLE V. COMPARISON OF DISTINGUISHABLE FEATURES OF VISUALIZATION LAYER

Feature	Our concept	Reducing dimensionality	Multi-dimensional data visualization (2D or 3D)
Flexible number of dimensions (any number of dimensions can be visualized without changing the overall appearance )	Yes	Yes (select most important 2 or 3 dimensions based on a selection algorithm)	Yes (example: scatter plot matrix, parallel coordinates plot, etc.)
Ability to compare behavior of its own dimensions (dimensions of a single data instance)	Yes (all dimensions will be visualized in same way)	No (not possible)	No (example: scatter plot matrix)
Ability to group by different dimensions over visualized dimensions (group by dim1 and dim2, but visualize dim1 and dim3)	Yes (SOM algorithm for grouping and different layer for visualizing)	No (not possible)	No (no separate layers for clustering and visualization)
Interactivity with the model	Yes	No	Yes

As explained in the later part of the section DATASET AND DIMENSIONS, there is no considerable variation with the values of answered table while there is a considerable variation with the values of abandon table. As per the below figure (Fig. 10), there is a high fluctuation in abandoned call counts over answered call counts. Our selection of abandoned table’s columns as dimensions was justified. This behavior was experienced mainly because of unavailability of extra contact center staff to cater a sudden demand.

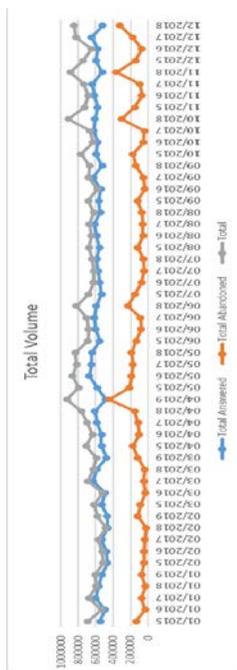


Fig. 10. Trend of Total Calls, Total Answered Calls and Total Abandoned Calls vs Time.

By combining two or more visualization and dimensionality reduction techniques, it is possible to have more features than sticking into a single visualization or dimensionality reduction technique. This can be concluded as the main finding of this research.

IX. CONCLUSION

Our approach showed improved results when predicting cause against traditional wallboards equipped with statistical analysis. SOM based approach shows better results when analyzing multidimensional data over traditional wallboards which shows charts with two or three axis. Results were highly depending on selection of dimensions correctly. In contrast with traditional wallboards, one wrong dimension selection may dilute the quality of overall prediction. Promising process for selecting dimensions will be a future work for this research.

ACKNOWLEDGMENT

I want to thank University of Colombo School of Computing, Colombo, Sri Lanka for guiding me to successfully complete my research project.

REFERENCES

- [1] C. Samaranyaka and S. Premaratne, “Enhancing call center operations through semantic voice analysis,” p. 10, 2018.
- [2] “24/7 CALL Center Solution: Business Purpose Call Center System with Asterisk PABX,” vol. 6, no. 9, p. 19, 2016.
- [3] Galkin, Popov, Pilyugin, and Grigorieva, “The Visualization Method Pipeline for the Application to Dynamic Data Analysis,” 2019.
- [4] M. I. Jordan and T. M. Mitchell, “Machine learning: Trends, perspectives, and prospects,” *Science*, vol. 349, no. 6245, pp. 255–260, Jul. 2015, doi: 10.1126/science.aaa8415.
- [5] N. Kour and D. N. K. Gondhi, “Recent trends & Innovations in Artificial Intelligence based Applications,” p. 6.
- [6] M. D. Assunção, R. N. Calheiros, S. Bianchi, M. A. S. Netto, and R. Buyya, “Big Data computing and clouds: Trends and future directions,” *Journal of Parallel and Distributed Computing*, vol. 79–80, pp. 3–15, May 2015, doi: 10.1016/j.jpdc.2014.08.003.
- [7] E. Olshannikova, A. Ometov, Y. Koucheryavy, and T. Olsson, “Visualizing Big Data with augmented and virtual reality: challenges and research agenda,” *Journal of Big Data*, vol. 2, no. 1, Dec. 2015, doi: 10.1186/s40537-015-0031-2.
- [8] S. P. Callahan, “Adaptive Visualization of Dynamic Unstructured Meshes,” p. 2.
- [9] D. Mashima, S. G. Kobourov, and Y. Hu, “Visualizing Dynamic Data with Maps,” p. 8.
- [10] E. R. Gansner, Y. Hu, and S. G. Kobourov, “GMap: Drawing Graphs as Maps,” arXiv:0907.2585 [cs], Jul. 2009, Accessed: Jul. 19, 2020. [Online]. Available: <http://arxiv.org/abs/0907.2585>.
- [11] M. Khan and S. S. Khan, “Data and information visualization methods, and interactive mechanisms: A survey,” *Int. J. Comput. Appl.*, vol. 34, no. 1, pp. 1–14, 2011.
- [12] C. Eze, J. R. C. Nurse, and J. Happa, “Using Visualizations to Enhance Users’ Understanding of App Activities on Android Devices,” p. 19.
- [13] D. P. Tegarden, “Business Information Visualization,” vol. 1, p. 38.
- [14] P. Hoffman and G. Grinstein, *Visualizations for high dimensional data mining-table visualizations*. Citeseer, 1997.
- [15] G. Grinstein, M. Trutschl, and U. Cvek, “High-dimensional visualizations,” in *Proceedings of the Visual Data Mining Workshop, KDD, 2001*, vol. 2, p. 120.
- [16] M. D. Assunção, R. N. Calheiros, S. Bianchi, M. A. S. Netto, and R. Buyya, “Big Data computing and clouds: Trends and future directions,” *J. Parallel Distrib. Comput.*, vol. 79–80, pp. 3–15, May 2015.

- [17] N. A. Qureshi et al., "Application of Principal Component Analysis (PCA) to Medical Data," *Indian J. Sci. Technol.*, vol. 10, no. 20, pp. 1–9, Feb. 2017.
- [18] L. N. Kova and O. S. T. Pánková, "Multidimensional clusters in RadViz," p. 6.
- [19] M. Schmuker et al., "SOMMER: self-organising maps for education and research," *J. Mol. Model.*, vol. 13, no. 1, pp. 225–228, Nov. 2006.
- [20] G. K. Uyanık and N. Güler, "A Study on Multiple Linear Regression Analysis," *Procedia - Soc. Behav. Sci.*, vol. 106, pp. 234–240, Dec. 2013.
- [21] F. Fechete, "Analysis of the Economic Performance of a Organization using Multiple Regression," p. 6, 2014.
- [22] E. Biglieri and K. Yao, "Some properties of singular value decomposition and their applications to digital signal processing," *Signal Process.*, vol. 18, no. 3, pp. 277–289, Nov. 1989.
- [23] Dr. G. Singh and A. Kaur, "Comparative Analysis of K-Means and Kohonen-SOM data mining algorithms based on student behaviors in sharing information on facebook," *Int. J. Eng. Comput. Sci.*, Apr. 2017.
- [24] H. S. Bhosale and D. P. Gadekar, "A Review Paper on Big Data and Hadoop," *Int. J. Sci. Res. Publ.*, vol. 4, no. 10, pp. 1–7, 2014.
- [25] M. Tory and T. Moller, "Human factors in visualization research," *IEEE Trans. Vis. Comput. Graph.*, vol. 10, no. 1, pp. 72–84, Jan. 2004.
- [26] S. Card, "Information Visualization," p. 37.