Clustering Approach for Congestion in Mobile Networks

J. Usha †, Ajay Kumar†† and A.D. Shaligram †††,

†Asst. Prof, Department of MCA, R.V.College of Engineering, Bangalore, India

††Director, JSPM's Jayawanth Institute of Computer Applications, Pune, India

†††H.O.D, Department of Electronic Science, University of Pune, India

Summary

The fast growing mobile networks demands faster and efficient approaches to handle the congestion in the network. This paper describes the clustering approaches like Partition around Medoids and Nearest Neighbour, used to determine the status of the base stations for mobile traffic load. This enables us to determine the number of channels to be borrowed from the co-channel cells to meet the increase in traffic load. Short-term traffic prediction is obtained by applying the radial basis function network. The traditional methods are not efficient in enhancing the performances because heavily-loaded cells sometimes cannot borrow unused nominal channels from their neighbouring cells which are idle or moderately loaded, as it may lead to cascading effect. The performance of the base stations can be increased if short-term traffic load can be predicted. The predicted results can then be used for channel borrowing from the idle or moderately loaded cells for long term use. These clustering approaches enhance the performance compared to traditional approaches. Simulation results also corroborate that the clustering methods enables the system to work with better performance than the traditional methods.

Key words:

Cluster, Partition around Medoids, Nearest Neighbour, radial basis function network, base station

1. Introduction

The Mobile networks have been extensively studied in recent years by both the computer and communication communities. However, most of the studies focus on the network layer protocols and few of them on the mobility aspect of the mobile user's behavior. It is desirable that the mobile load on a base station is anticipated. Traditional resource management passively tracks the network [6] and collects the mobile data. Existing channel allocation and borrowing algorithms are investigated and observations are made. There is a growing tendency towards optimal mobile network resource management. This affects other aspects such as capacity planning, operations and allocation to meet the increasing demand on planning management and operations of the mobile network. The operation and planning of a mobile network requires an appropriate model for predicting the mobile traffic load. Load forecasting and their accuracy plays a key role in helping the base stations and the Mobile switching centres to make important decisions on capacity, channel allocation, load balancing, switching control, network reconfiguration, and infrastructure development. Dynamic load balancing is commonly used in processor scheduling in distributed systems to achieve better processor utilization [1]. In a dynamic load balancing scheme [1,2,3,4,5,6,7] for the channel assignment problem [14] in a cellular mobile environment, unused channels from under loaded cells are migrated to an overloaded one by borrowing a fixed number of channels from cold cells to a hot one using a channel borrowing algorithm. Initially, each cell is allocated a set of channels, each to be assigned on demand to a user in the cell. A cell is hot[1] if the degree of coldness of a cell defined as the ratio of the number of available channels to the total number of channels for that cell is less than or equal to some threshold value, other wise the cell is cold. Here a Markov model for an individual cell has been developed where the cell state is determined by the number of occupied channels in the cell. The number of channels in a compact pattern is given by $N = i^2+ij+j^2$ which determines the number of channel sets to be formed. The performance of the load balance for channel assignment using selective strategy is compared with the fixed channel assignment, simple borrowing, directed retry and channel borrowing with locking. This strategy shows significant improvement in the performance over the other methods.

An alternate approach using simulated annealing [8,15] has been proposed. Fixed schemes perform well under heavy traffic conditions but if the number of calls exceeds the channel set for the cell, the excess calls are blocked until channels become available. To address this problem, channels are borrowed from other cells. Borrowing strategies can be simple, hybrid and channel ordering [9,10]. In the Simple Borrowing scheme, channels are borrowed provided existing calls are not affected. The borrowed channel is locked in those co-channel cells of the lender which are non-co-channels of the borrower. This approach shows poor performance during heavy traffic conditions, as the channels are locked. In the Hybrid Borrowing scheme, the fixed channel set assigned to a cell is divided into local use only and lending channels on demand. The number of channels in each group is determined apriori depending on the history of

Manuscript received February 20, 2010
Manuscript revised February 27, 2010

the traffic conditions. Channel ordering technique extends the hybrid scheme, where the number of channels in the hybrid borrowing scheme can vary dynamically depending on the traffic conditions. Each channel is ordered such that the first channel has the highest priority of being locally used and the last channel has the highest priority of being borrowed. The ordering may change depending on the traffic pattern. A released higher order channel is relocated to an ongoing call in a lower order channel and hence reduces locking of borrowable channels.

In the next section the general mobile network architecture is described.

2. The Mobile Network Architecture

A given geographical area consists of hexagonal cells each served by a base station (BS). The BS is a part of the wireless system infrastructure that controls one or multiple cell sites and radio signals [13]. The BS thus reduces the load on the switch. BS performs radio signal management functions for base transceiver stations, managing functions such as frequency assignments and handoff. This is included as a part of the switch and the mobile station or the user (MS) communicate through wireless links and form localized wireless networks. The general architecture for the mobile network is shown in Fig 1.

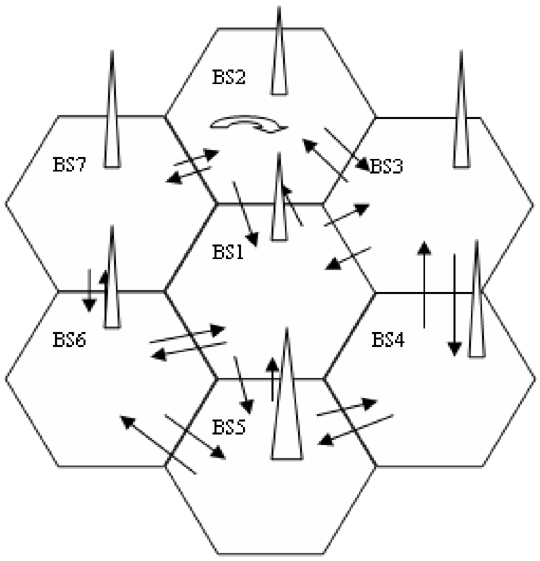

Fig 1: The Mobile Network Architecture

A group of cells are in turn served by a mobile switching center (MSC). Each cell is allocated a fixed set of C channels[12] according to the compact pattern based fixed assigned scheme. Initially, we consider Mobile stations can be stationary or roaming. Traffic can move within the base station or between base stations. Like within BS1, BETWEEN BS1 – BS2, BS1- BS3, BS2-BS3, BS2-BS1, WITHIN BS2, BS3-BS1, BS3-BS2, WITHIN BS3.Traffic can move forward or reverse. Blocking describes the situation when a user attempts to make a call and is not able to reach a dialled subscriber due to lack of resources. Generally it is assumed that a user will generate about 25mErl of traffic[6] during the busy hour and that the average network usage is about 120 seconds. This may vary from network to network. Some networks use average 35mErl and 90s.

In the next section, the problem is defined.

3. Problem Definition

When a mobile user intends to communicate to another user or a base station, the user must first obtain a channel from one of the base station that hears it. If a channel is available, it is granted to the user, otherwise the call is blocked. This call is called a new call. The subscriber releases the channel either when the user completes the call or when he moves to another cell before the call is completed. The process of moving from one cell to another when a call is in progress is called a handoff call. While performing handoff, the mobile unit requires that the base station in the cell that it moves into will allocate the channel. If no channel is available in the new cell, the handoff call is blocked. Retry calls[20] is a call made by a subscriber who was unsuccessful in getting the call established. We use the clustering approach [20-25] to form clusters based on the traffic load. The approaches used are Partition around Medoids (PAM) and Nearest Neighbour clustering. These are conventional clustering approaches. They are used to classify the base stations into heavily loaded, moderately loaded and idle base stations. In order to determine from which neighbouring cell the channels need to be borrowed, radial basis function a neural network based approach is adopted. This in turn helps in handling congestion issues in a more efficient manner. This is compared to the fixed channel allocation scheme. Clustering is a discovery process, the differences between the datasets can be discovered in order to separate them into different groups and similarity between data sets can be used to group similar data together.

Clustering helps in quality control by detecting experimental artifacts/bad hybridizations. It also helps check whether samples are grouped according to known categories. It also enables one to identify new classes of biological samples (e.g. tumor subtypes).

Generally, cluster analysis is based on two factors namely the distance measure and the cluster algorithm. The distance measure is the quantification of similarity or dissimilarity of objects. The Cluster algorithm is a procedure to group objects in to Small, based on within-clusters distances and Large, based on between-cluster

distances. The choice of distance measure to be used is based on the application area. Correlation distance d_c measures trends/relative differences for applications to detect similarities. $d_c(x, y) = d_c(ax+b, y)$ if $a > 0$.
Euclidean and Manhattan distance both measure absolute differences between vectors. Manhattan distance is more robust against outliers. Standardization may be applied to the observations: by subtract the mean and dividing by standard deviation. After standardization, Euclidean and correlation distances are equivalent. As one particular algorithm for clustering with a restricted function space, we have the Nearest Neighbour clustering. This algorithm is seen as a general baseline algorithm to minimize arbitrary clustering objective functions. In essence, it works with a class of functions which are constant on local neighborhoods. This function class is small. We allow the function class to grow slowly with n by allowing the neighborhood to become smaller and smaller. In the limit for n-> ∞, we can then approximate any clustering on the underlying data space up to arbitrary precision. Nearest neighbour clustering selects random seed points from the data and only considers candidate clustering which are constant. Among those clustering, NNC chooses the best one according to a given clustering objective function. This method can be used to choose the unused channels from the co-channel cells. The side effect of the nearest clustering approach is that problems in clustering are NP hard and thus inherently difficult to solve. In the next section, the PAM clustering approach is described.

4.0 Partitioning around Medoids (PAM): Partitioning around medoids (PAM) generalizes the idea of being based on Euclidean distance and can be used with any distance measure *d* (objects x_i need not be vectors). The cluster centers/prototypes are required to be observations. They try to minimize the sum of distances of the objects to their cluster centers,

$$\sum_{i=1}^{n} d(x_i, m_{j(i)}),$$

using an iterative procedure analogous to the one in *K*-means clustering.
To choose *K* (the number of clusters), many heuristic approaches try to compare the quality of clustering results for different values of *K*. The problem can be better addressed in model-based clustering, where each cluster represents a probability distribution, and a likelihood-based framework can be used.

Limitations of clustering

Clustering has been very popular in the analysis of micro array data. However, many typical questions in micro array studies (e.g. identifying expression signatures of tumor types) are better addressed by other methods (classification, statistical tests). Clustering algorithms are easy to apply, and they are useful for exploratory analysis. But it is difficult to assess the validity/significance of the results. Even "random" data with no structure can yield clusters or exhibit interesting looking patterns [26]. In the next section the Nearest neighbour clustering approach is described.

4.1 Nearest Neighbour Clustering Method

is a method of calculating distance between clusters in hierarchical cluster analysis. The linkage function specifying the distance between two clusters is computed as the minimal object to- object distance d (x_i, y_j), where objects x_i belong to the first cluster, objects y_j belong to the second cluster. In other words, the distance between two clusters is computed as the distance between the two closest objects in the two clusters[11].

4.1.1 Nearest Neighbour clustering Algorithm

1. Set i = 1 and k = 1. Assign pattern x_1 to cluster c_1.
2. Set i = i + 1. Find nearest neighbour of x_i among the patterns already assigned to clusters. Let d_m denote the distance from x_i to its nearest neighbour. Suppose the nearest neighbour is in cluster m.
3. If d_m greater than or equal to t then assign x_i to c_m where t is the threshold specified by the user. Otherwise set k = k+1 and assign x_i to a new cluster c_k.
4. If every pattern has been considered then stop else go to step 2.
5. The nearest neighbour clustering method is a method of calculating distance between clusters in hierarchical cluster analysis. The linkage function specifying the distance between two clusters is computed as the minimal object-to-object distance neighbour, where objects x_i belong to the first cluster, objects y_j belong to the second cluster. In other words, the distance between two clusters is computed as the distance between the two closest objects in the two clusters.
6. Mathematically the linkage function - the distance D(X,Y) between clusters X and Y- is described by the following expression :

$$D(X,Y) = \min_{x \in X\, y \in Y} d(x,y)$$

Where

d(x,y) is the distance between objects x and y; X and Y are two sets of objects (clusters)

This algorithm initializes the cluster centroids with hierarchical clustering and hence gives deterministic outcomes.

Table 1 Comparison of PAM and Nearest Neighbor

No. of Clusters	Time taken(Secs)	
	PAM	Nearest Neighbour
2	0.045	0.5
4	0.067	0.7
6	0.075	0.8
8	0.087	0.9

From the table, we infer that as the number of clusters increases, the time taken for PAM increases but is less when compared to Nearest Neighbour algorithm. The performance evaluation of these algorithms is based on time to identify the clusters. In the next section, the radial basis function network approach is described.

4.2 Radial Basis Function Network

Step 1: Initially, random data points are considered to represent the base stations and the density of the traffic.
Step 2: The Algorithm based on RBF is applied once.
The different Base Stations are clustered based on proximity, capacity, density of the traffic as idle, moderately or heavily loaded.
Step 3: Apply neural network, predict the number of co-channels required to be borrowed.
Step 4: Since co-channel borrowing depends on neighboring BS. There is a tendency of the co-channel cells being / becoming equally loaded. This is obtained from the neighboring BS guard band / threshold. μ in the RBF is determined. If μ is inversely proportional to the guard band, go to the farther BS so that long time, channel from the co-channel can be borrowed.
Thus the model is created.

Co-channel borrowing scheme is optimized with regard to available resources, current load and the neighboring BS. Load balancing using Backpropagation method is discussed in [19]. Radial basis function networks can be used for prediction applications [16-18]. Here RBF is used to determine from which co-channel cells the borrowing may be done. This is done based on μ being decided by the guard band / threshold. Threshold here implies the maximum amount of co-channel cells that may be borrowed say 95%.

Case 1: Clusters belong to the same vendor
Clusters belong to the same vendor, whenever borrowing is within the cluster. Consider that base station X needs 4% of capacity of co-channels to be borrowed. Also the guard bands of the neighbour BS X1, X2, X3 are 4%, 3.9% and 6% respectively. Now the RBF will also make the μ_1 of X_1 proportional to 4% and μ_2 of X_2 proportional to 3.9% etc.
Since μ is inversely proportional to GB though X1, X2 and X3 are at the same distance from X (if K nearest neighbour rule is applied) But the RBF with variable μ will make X3 as the nearest as compared to X1 and X2 as it takes into consideration as the nearest to X.

Case 2: Clusters belonging to different vendors

Clusters belong to different vendors whenever borrowing is outside the cluster. Then other issues will come up like mutual understanding between the vendors. Based on the point of access, the channel borrowing may be done.

4.2.1 Effect of co-channel borrowing in the cluster

1. Cascading effect- Say X is loaded; borrowing 4% from X_1 leads to X_1 being overloaded, which is true with its internal dynamics(ie with its own calls) may leads to borrowing from X_2 etc.. Say X borrows only from X_3 which has an additional 2% capacity to spare and then this cascading effect is controlled.
2. Co-channel borrowing among nearest neighbors is generally the trend with all conventional approaches like K-nearest, PAM whereas with RBF, based on traffic density ie Guard Band and the cluster it belongs to. (ie distance is adaptive)

Generally, distance from nearest neighbour is fixed i.e. geometrical distance (GD).

Here in RBF, $\mu_A \propto GD$ and $\mu_B \propto \frac{1}{GB}$. As μ_B increases, GB decreases.

$\mu = \mu_A + \mu_B$. Since μ_A is fixed component, as μ increases, μ_B increases ie if GB decreases, μ_B increases.

This implies if the nearest BS are overloaded, TD is more, GB is less; the μ of that is very large and they appear to be farthest and the BS say X_6 in the above example appear to be nearest.

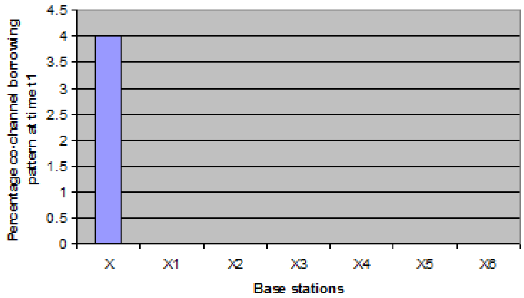

Fig 2. Trend graphs of co-channel borrowing pattern at the time t1.

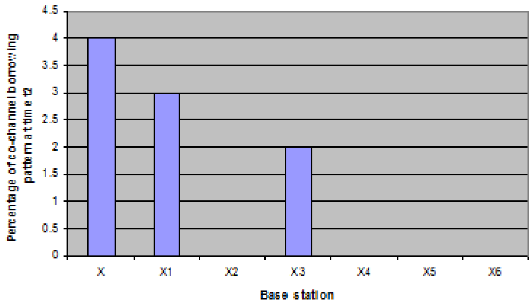

Fig 3: Trend graph of co-channel borrowing pattern at time t2

In Fig 3, the trend graph at time t2 and the variation in traffic density with increase of X1 and X3, X and X2 remaining same is depicted. Here X1 lends some channels to X which cannot be immediately returned back i.e returned only after call termination, the borrowed co-channels can be taken back from X1) else this lending affects X1. Total co-channel borrowing in the cluster is 9%.

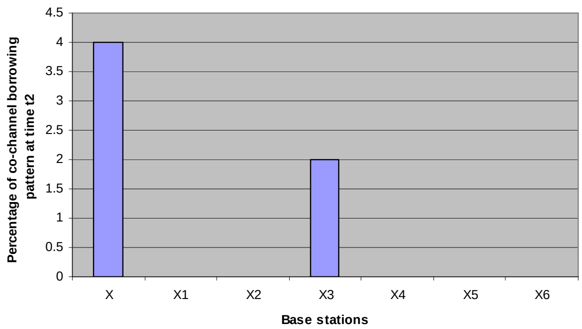

Fig 4. Trends graphs of co-channel borrowings in a cluster of base stations.

In fig 4, we find that with X borrowing from X6 using our RBF, the total co-channel borrowing is 6% in the cluster which is less compared to borrowing from other co-channel cells. Hence by using this approach we will be borrowing less number of channels as compared to traditional approaches.

5. Conclusions

In this paper, we considered the situation where mobile cells are clustered based on the density of the mobile traffic load. They are identified as idle, moderately and heavily loaded cluster cells. We have used conventional clustering techniques such as nearest neighbour clustering and PAM clustering method to identify the clusters. After identification of the clusters, we adopted a congestion control scheme using radial basis network function where heavily loaded clusters are surrounded by moderately loaded and idle clusters. This helps us to identify cells from which channels can be borrowed. Short term prediction of the traffic load using the radial basis function can be adopted to predict the traffic load. After detection, we determine how many co-channels need to be borrowed. This is borrowed from the cells which are considered to be nearest using radial basis function network. This is compared to the Fixed Channel Allocation method. Simulations demonstrate that our adoption of the clustering approach can deliver better performance than Fixed Channel Allocation method.

References

[1] Sajal K. Das, Sanjoy K. Sen and Rajeev Jayaram,"A Novel Load Balancing Scheme for the Tele-Traffic Hot Spot Problem in Cellular Networks.

[2] Sajal K. Das and Sanjoy K. Sen and Rajeev Jayaram," A Structured Channel Borrowing Scheme for Dynamic Load Balancing in Cellular Networks,International Conference on Distributed Computing Systems,1997

[3] Guohong Cao,"Integrating Distributed Channel Allocation and Adaptive Handoff Management for QoS-Sensitive Cellular Networks,Wireless Networks,2003, vol 9

[4] Sajal K. Das, Sanjoy K. Sen and Rajeev Jayaram,"A Dynamic Load Balancing Strategy for Channel Assignment Using Selective Borrowing in Cellular Mobile Environment,wireless networks, 1996,pp 1-25.

[5] S.K. Das, S.K. Sen, R. Jayaram,"A structured channel borrowing scheme for dynamic load balancing in cellular networks," 17th IEEE International Conference on Distributed Computing Systems (ICDCS '97),Baltimore, May 27-May 30,1997,pp 116

[6] Ajay R Mishra, "Fundamentals of Cellular Network Planning and Optimization-2G/2.5G/3G… Evolution to 4G", Wiley, 2004, England.

[7] M.H. Willebeek-LeMair and A.P. Reeves,"Strategies for dynamic load balancing on highly parallel computers", IEEE Trans. On parallel and distributed systems, vol 4, No. 9, Sept 1993.

[8] M Duque-Anton, D. Knuz and B. Ruber, " Channel Assignment for Cellular Radio using Simulated Annealing", IEEE Trans. Vehicular Technology,vol VT-22, pp 218-222, Nov 1973.
[9] S.M. Elnoubi, R. Singh, S.C. Gupta, " A New Frequency Channel Assignment in High Capacity Mobile Communication Systems", IEEE trans, Veh, technology, vol VT-31, no.3, August 1982
[10] S.Tekinay and B. Jabbari, " Handover and Channel Assignment in Mobile Cellular network", IEEE Communication magazine, Nov 1991.
[11] Bubeck S. U. Von Luxburg. "Nearest Neighbour Clustering: A Baseline Method For Consistent Clustering with Arbitrary Objective Functions. Journal of Machine Learning Research, 2009.
[12] Katzela I, Naghshineh M (1996) Channel assignment schemes for cellular mobile telecommunication systems: a comprehensive survey. IEEE Personal Commun 3(3):10–30
[13] Lee WCY (1989) Mobile cellular telecommunications systems. McGraw-Hill, New York
[14] Tamura H, Sengoku M, Shinoda S, Abe T (1991) Channel assignment problem in a cellular system and a new coloring problem of networks. IEICE Trans Commun Electr Inf Syst 74(10):2983–2989
[15] Kunz D (1991) Channel assignment for cellular radio using neural networks. IEEE Trans Vehicular Technology 40(1):188–193
[16] Leung H, Lo T, Wang S (2001) Prediction of noisy chaotic time series using an optimal radial basis function neural network.IEEE Trans Neural Netw 12(5):1163–1172
[17] Chang FJ, Liang JM, Chen YC (2001) Flood forecasting using radial basis function neural networks. IEEE Trans Syst Man Cybern Part C 31(4):530–535
[18] Hansen JV, Nelson RD," Neural networks and traditional time series methods: a synergistic combination in state economic forecasts". IEEE Trans Neural Netw 8(4):863–873,1997.
[19] J. Usha, Ajay Kumar and A.D. Shaligram," Load balancing using back propagation in mobile networks", to appear in CIIT International Journal of Wireless networks, Coimbatore, India, Feb 2010.
[20] J.Han and M. Kamber," Data Mining Concepts and Techniques" 2nd edition
[21] Margaret H. Dunham and Sridhar,"Data Mining, Introduction and Advanced Topics", Pearson Education, ISBN 81-7758-785-4
[22] Arun K Pujari "Data Mining Techniques", Universities Press (India) Limited,ISBN 81-7371-380-4
[24] Pieter Adriaans & Dolf Zantinge "Data Mining,",Pearson Education Asia, ISBN 81-7808-425-2. Addison Wesley Longman (Singapore).
[25] Michael J. A. Berry and Gordon S. Linoff "Data Mining Techniques for Marketing, Sales and Customer Relationship Management, , Wiley-Dreamtech India Pvt. Ltd., ISBN 81-265-0517-6.
[26] Lamiaa Fattouh Ibrahim, Manal Hamed Al Harbi, "Using Clustering Technique M-PAM in Mobile Network Planning",12th WSEAS International Conference on Computers, Greece, July 23-25,2008, 868-873.

J. Usha has received her B.Sc in 1994 and MCA in 1998 from Bangalore University, Bangalore, India. She is currently pursuing her research in the area of Mobile Computing in Department of Computer Science, University of Pune, Pune. Her major field of study is mobile computing and computer networks.
She is working as Asst Professor, Department of Master of Computer Applications, R.V. College of Engineering, Bangalore, India. She has over eleven years of teaching / research experience. She has published many papers in National/ International conferences and Journals. She is a life member of ISTE and CSI.

Dr. Ajay Kumar received his M.Sc. Applied Science (Computer), in 1992 and Ph.D in 1995 from Marathwada University, Aurangabad, India. His major of study include data communication, networking and software engineering.
He is working as Director, at Jayawant Institute of Computer Applications, Pune, India. He has over 20 yrs of industry/ teaching experience. He has authored 4 undergraduate books in computer science. He has published many papers in International and National Journals and Conferences. Dr Ajay Kumar is a life member of CSI.

Dr. A. D. Shaligram received his B.Sc in 1979, M.Sc in 1981 and Ph.D in 1986 from University of Pune, India. He became a member of IEEE in 1993. His major areas of study include Optoelectronics, Fiber optic and Optical Waveguide sensors. PC / Microcontroller based Instrumentation, Simulation software development, Biomedical Instrumentation and sensors. He is working as H.O.D, Electronics Science Department, University of Pune. He has 24 yrs of teaching / industry experience. He has authored 20 books in electronics science. He has published many papers in National / International Conferences and Journals. He is a member of IAPT and IPA.