International Innovative Research Journal of Engineering and Technology

ISSN NO: 2456-1983

International Conference on Emerging Innovation in Engineering and Technology

ICEIET-2017

Reliable And Ascendable Content Based Image Retrieval Aproach In Peer To Peer Networks

[1]S.BRINDHA, [2]N.KAVITHA, [3]R.MANJULA, [4]Ms.K.VISHNUPRIY

MRK Institute of Technology Kattumannarkoil

brindhapriyanga@gmail.com, kavithakavi686@gmail.com , manjulacse96@gmail.com

ABSTRACT:

The peer to peer network is one of the major networks in now days, due to the accessing and sharing of data with minimum end to end delay. There are different types of image retrieval is present, so this project concentrates on the analysis of the different technology use for the image retrieval. Image retrieval is one of the major tasks from the internet, if we want any particular image we don't get that exact image from the large variety of set, so this project introduce the efficient method for image retrieval. In this project, we present experiments with efficient Content Based Image Retrieval in a P2P environment, thus a P2P-CBIR system.We propose a CBIR scheme based on codebook generation . A dynamic codebook updating method by optimizing the mutual information between the resultant codebook and relevance information, and the workload balance among nodes that manage different codewords.Experimental results will prove the effectiveness of the approach.

I. Introduction:

PEER-TO-PEER (P2P) networks, which are formed by equally privileged nodes connecting to each other in a self-organizing way, have been one of the most important architectures for data sharing. Popular P2P file-sharing net- works such as eDonkey[1] count millions of users [1] and tens of millions of files. Unlike webpages which mainly consist of textual documents such as news, blog articles or forum posts, multimedia files play a dominant role in most P2P networks [2]. The ever-growing amount of multimedia data and computational power on P2P networks exposes both the need and potential for large scale multimedia retrieval applications such as content-based image sharing, and copyright infringement detection.While P2P networks are well known for their efficiency, scalability and robustness on file sharing, providing extended search functionality such as content-based image retrieval (CBIR) faces the following challenges: 1) in contrast to centralized environments, data in P2P networks is distributed among different nodes, thus a CBIR algorithm needs to index and search for images in a distributed manner; 2) unlike distributed servers/clouds, nodes in P2P networks have limited network bandwidth and computational power, thus the algorithm should keep the network cost low and the workload among nodes balanced; 3) as P2P networks are under constant churn, where nodes join/leave and files publish to/remove from the network, the index needs to be updated dynamically to adapt to such changes.

To support content indexing and avoid message flooding, structured overlay networks such as Distributed Hash Tables (DHTs) [3], [4] are often implemented on top of a physical net- work. By organizing the nodes in a structured way, messages can be efficiently routed between any pair of nodes, and the index integrity can be maintained during network churn. For the CBIR functionality, most of the existing systems adopt a global feature approach: an image is represented as a high- dimensional feature vector (e.g., color histogram), and the similarity between files is measured using the distance between two feature vectors [5], [6], [7]. Usually, the feature vectors are indexed by a distributed high-dimensional index or Locality Sensitive Hashing (LSH) over the DHT overlay. However, due to the limitation known as "curse of dimensionality", the majority of these solutions have high net- work costs or serious workload balance issue among nodes when the dimensionality of feature vectors is high.On the other hand, the bag-of-visual-words (BoVW) model has been successfully utilized for large scale image retrieval [8]. In the BoVW model, each image is represented with a bag of local features, which mimics the bag-of-words (BoW) model where each document is a collection of unor- dered words.[2] Generally, to employ the BoVW model, the following three steps are required [9], [10]: Firstly, a number of local regions (through image segmentation or uniform image partitioning) or keypoints (through keypoint

ISSN NO: 2456-1983

detec- tion algorithms such as Hessian-Affine detector [11]) will be identified from an image and each region or keypoint will be represented with a high dimensional descriptor. In our experiments, the widely used Scale-Invariant Feature Trans- form (SIFT) descriptor [12] is employed.

II. Existing System:

The existing systems adopt a global feature approach: an image is represented as a high dimensional feature vector (e.g., color histogram), and the similarity between files is measured using the distance between two feature vectors. Usually, the feature vectors are indexed by a distributed high-dimensional index or Locality Sensitive Hashing (LSH) over the DHT overlay In contrast to centralized environments, data in P2P networks is distributed among different nodes, thus a CBIR algorithm needs to index and search for images in a distributed manner. P2P networks are under constant churn, where nodes join/leave and files publish to/remove from the network, the index needs to be updated dynamically to adapt to such changes. P2P networks are under constant churn, where nodes join/leave and files publish to/remove from the network, the index needs to be updated dynamically to adapt to such changes. Dexing and Locality-Sensitive Hashing. The high-dimensional indexing based approaches store the feature vectors in a data structure, usually a tree or a graph, to achieve effective search space pruning during retrieval. In structured P2P networks, the high-dimensional index is defined in a distributed way over the P2P overlay, dexing and Locality-Sensitive Hashing.

Disadvantage Of Existing System:

Even in a centralized environment, the performance of high-dimensional indexing suffers from the well-known "curse of dimensionality".Even when one can update the hash functions with changing data, implementing it over the DHTs is very challenging. As the data is stored among nodes of corresponding hash ID, a 1-bit change of the hash function output will result in large portion of (if not all) data being assigned to a different node, causing heavy network traffic.

III. Proposed System:

In this paper, we present a novel method to dynamically generate and update a global codebook, which considers both the discriminability and workload balance.While processing queries, each node collects the relevance information and workload data. With the relevance information, we maximize the information provided by the codebook about the retrieval results, thus minimizing the information loss incurred by quantization. With workload data, we aim to achieve a fair workload among nodes, thus avoiding overloading or underloading nodes. Based on these two criteria, the codebook partitioning is updated routinely by splitting/merging codewords, thus allowing the codebook to grow/shrink in accordance to the data distribution. To minimize the cost of codebook updating, the decision whether a codeword should be split/merged is taken by its managing node individually. Finally, the updates are synchronized across the network at the end of each iteration. As a result, the discriminability and workload balance is optimized continuouslywith the churn of the P2P network.

Advantage Of Proposed System:

It is the first study to investigate scalable CBIR with the BoVW model in P2P networks.A novel objective function for codebook optimization in a P2P environment is proposed, which considers both the relevance information and the workload balance simultaneously.A distributed codebook updating algorithm based on splitting/merging of individual codewords is proposed, which optimizes the objective function with low updating cost

SYSTEM ARCHITECTURE:

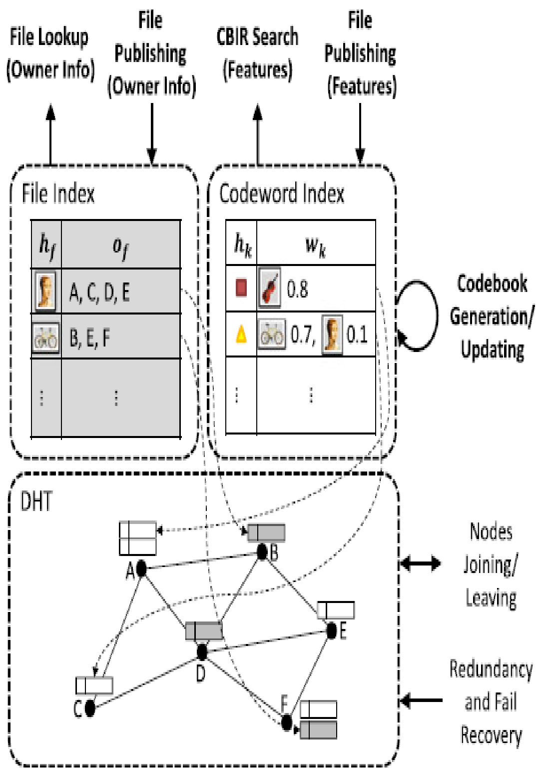
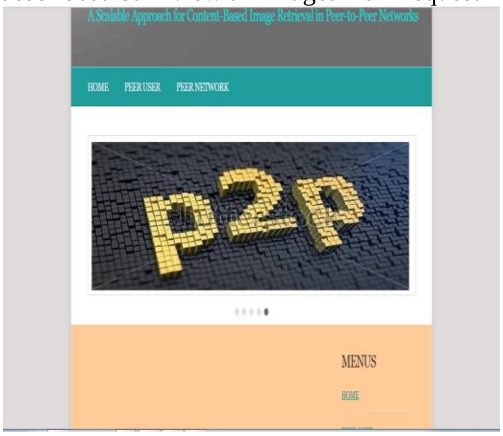


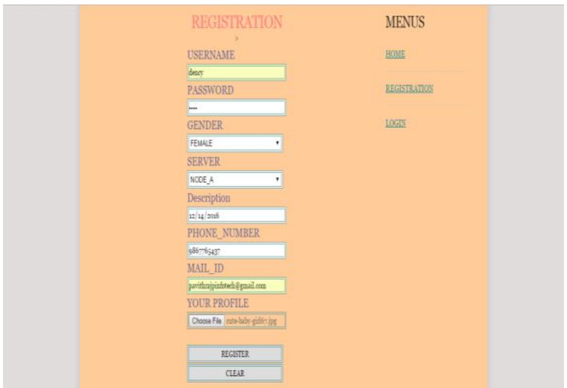
Figure 1. Architecture

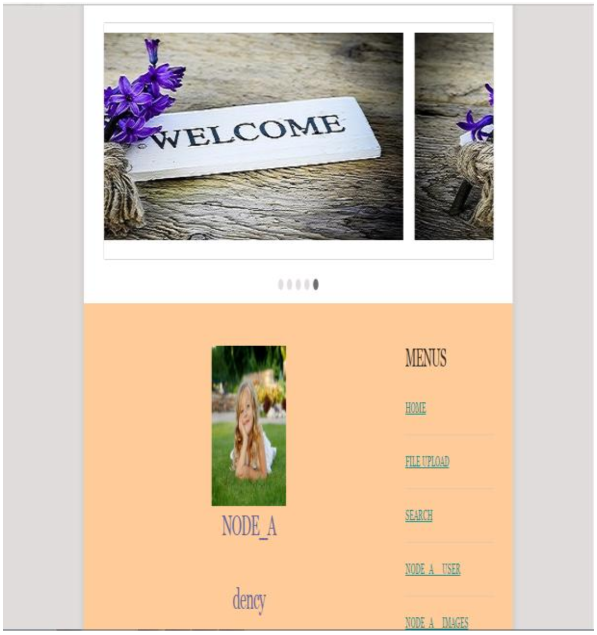
ISSN NO: 2456-1983

IV. Modules :CBIR Search:

The CBIR search is essentially an inverted index lookup in the codeword index. a user on node C submits a CBIR query with an example image (red bike), which will be answered in three steps: Firstly, the BoVW codewords will be extracted on node C locally. Secondly, the codewords will be looked up in the codeword index with GET(k), where k is the codeword ID (solid arrows). The postings with their corresponding file IDs will be returned by the respective nodes (E and A), which will be used as similarity measurement to produce the ranked results for the user. Finally, the owner information of relevant images m can be obtained by the file lookup process described. Can view all images from request Nodes.

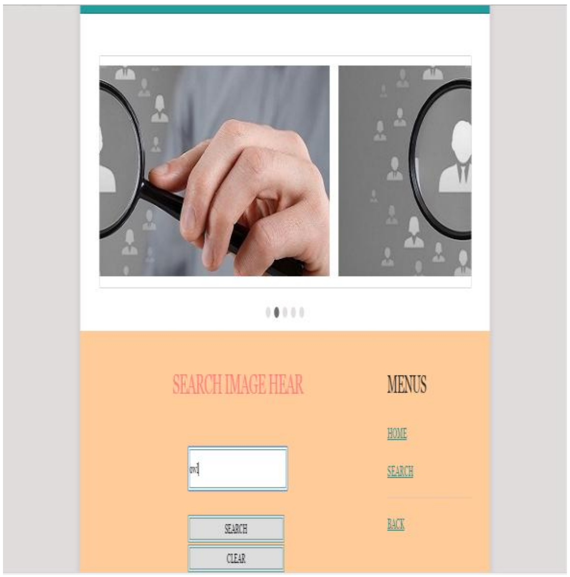


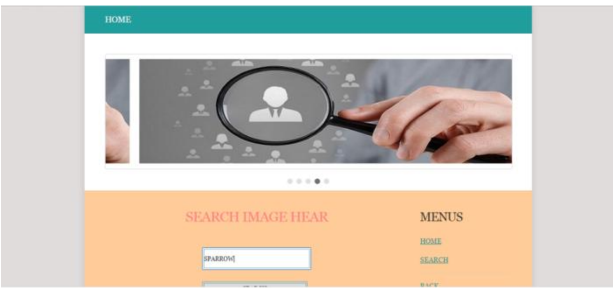


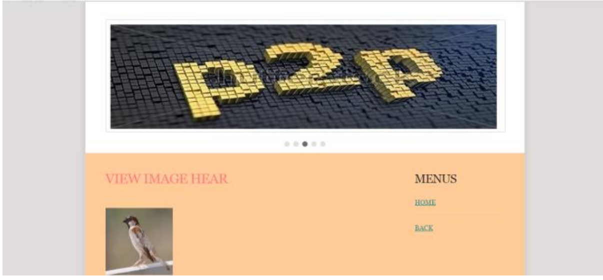


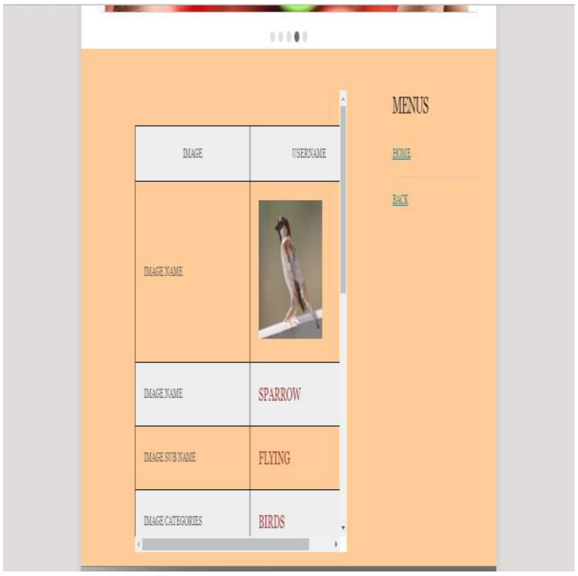
ISSN NO: 2456-1983

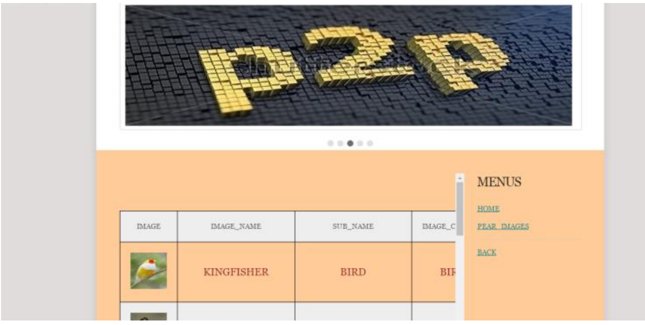
Bovw Model Retrieval Process: BoVW based representation for the query, retrieving the postings via DHT lookup, and measuring the similarity between the query and candidate images. In large scale BoW based retrieval systems, index pruning has been used to reduce the retrieval cost. Its basic idea is to identify and discard the postings which are not likely to contribute to top





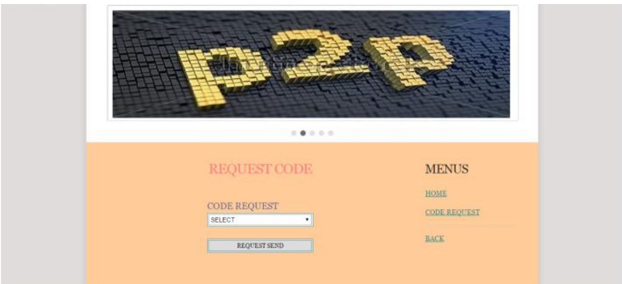


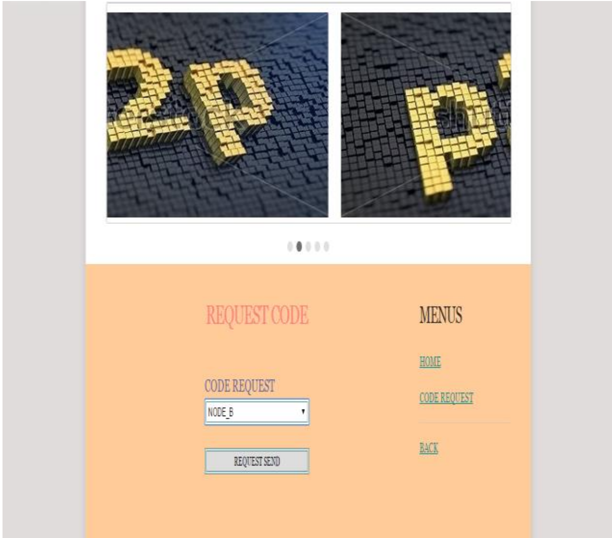


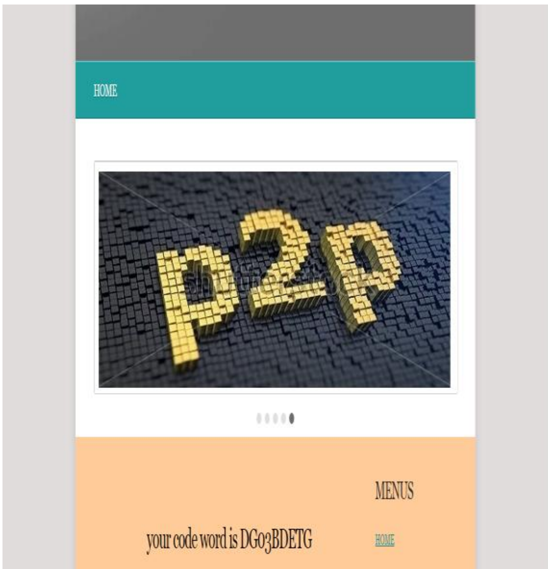


Codebook Generation And Updating

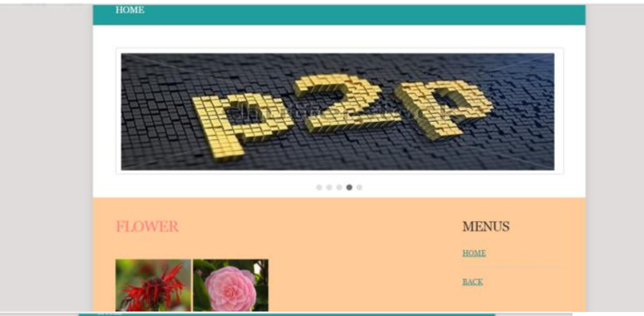
During an updating iteration, each codeword node pk decides whether its codeword k should be split/merged/unchanged based on the relevance information collected from past queries, and the current workload. After each iteration, the centroid coordinates and the codeword statistics needed for similarity measurement (e.g., document frequencies) will be broadcasted throughout the network, so that all the nodes in the network can have the same codebook. The iterative process runs continuously in order to maintain an updated codebook during data churn. The frequency of update iterations is determined .To split the codeword k into n codewords, pk randomly selects n1 neighboring nodes as new codeword nodes and sends the centroid coordinates to them. Once all the new centroids register themselves as codeword nodes, the descriptor associations of selected nearby partitions will be updated respectively similar to the file posting process.

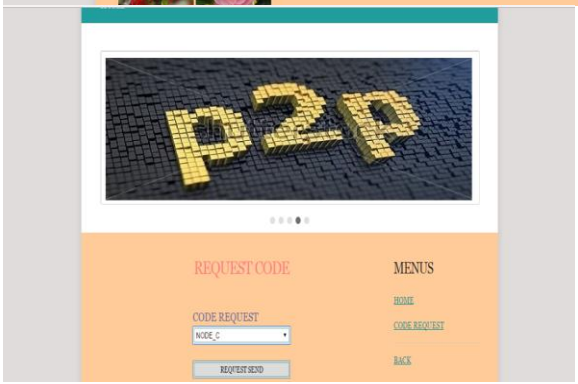


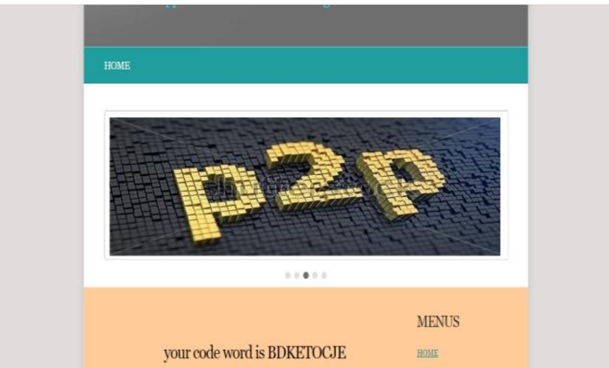


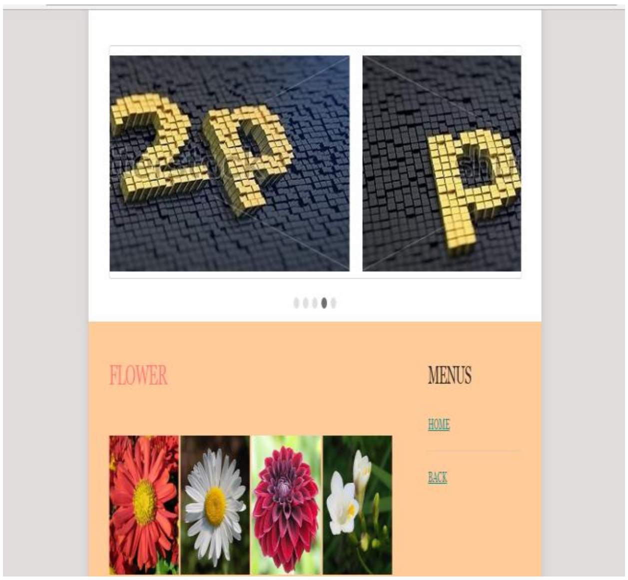


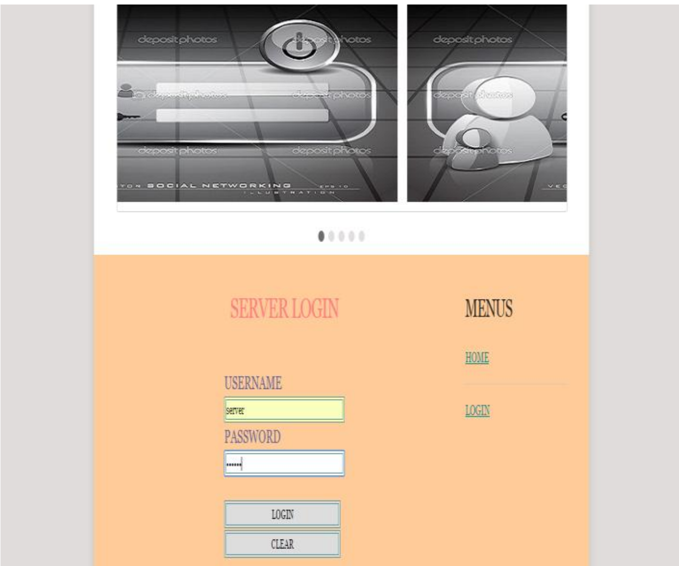
ISSN NO: 2456-1983



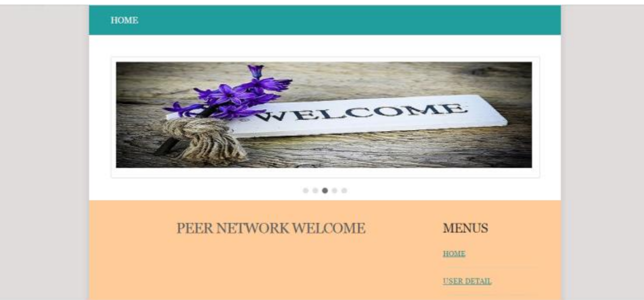


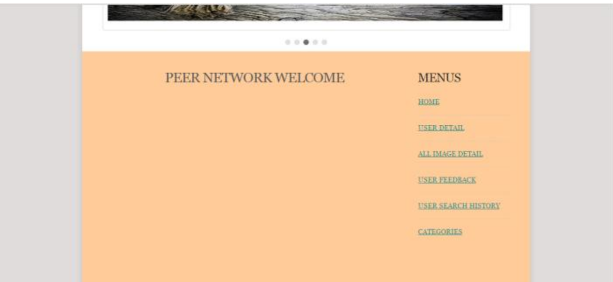


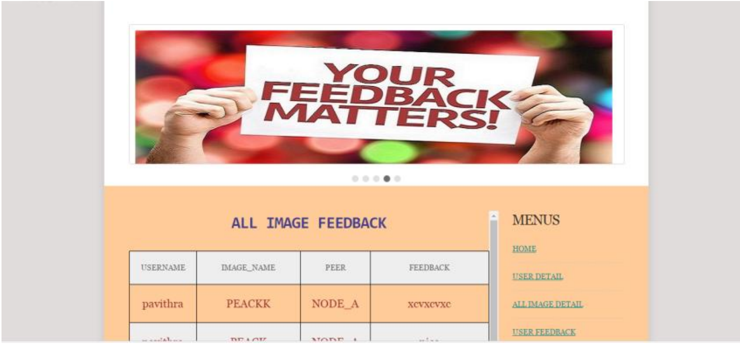


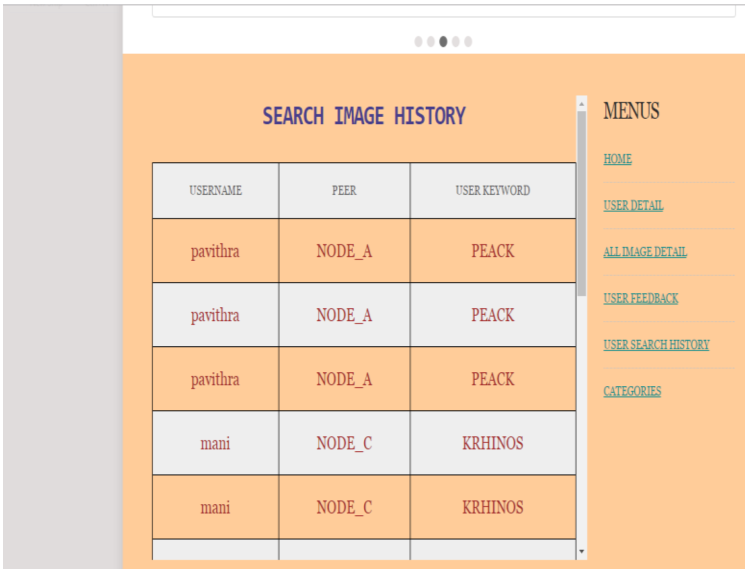


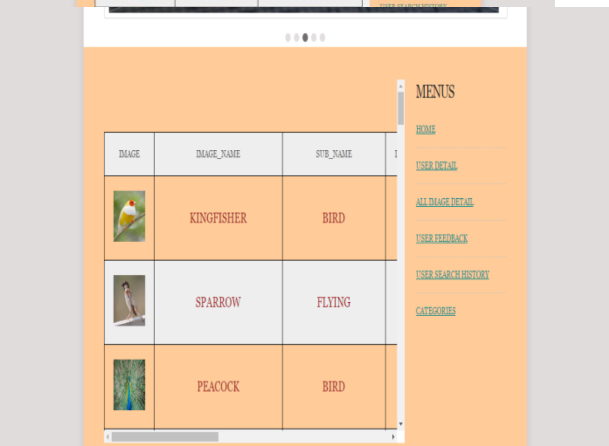
ISSN NO: 2456-1983

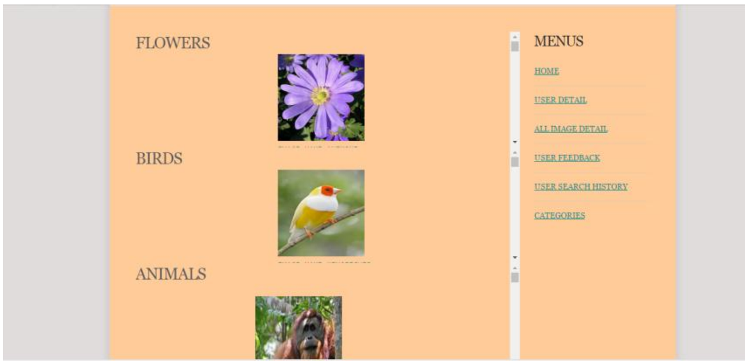


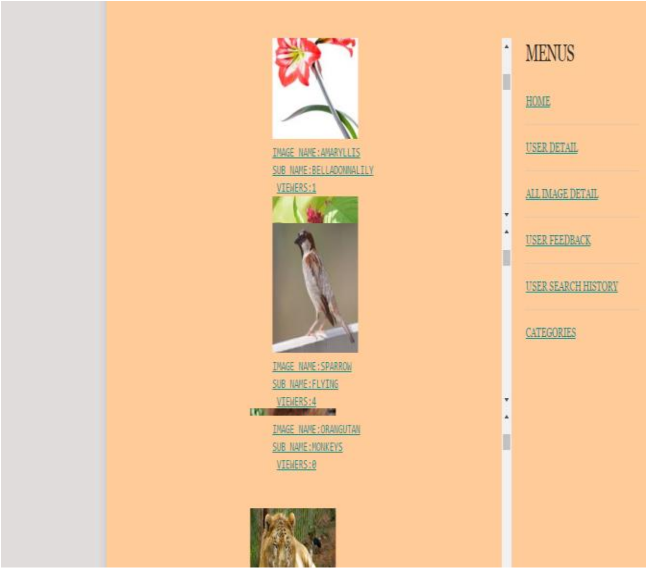











MENUS
HOME
USER DETAIL
ALL IMAGE DETAIL
USER FEEDBACK
USER SEARCH HISTORY
CATEGORIES
IMAGE_NAME:AMARYLLIS
SUB_NAME:BELLADONNALILY
VIEWERS:1
IMAGE_NAME:SPARROW
SUB_NAME:FLYING
VIEWERS:4
IMAGE_NAME:ORANGUTAN
SUB_NAME:MONKEYS
VIEWERS:0


MENUS
HOME
USER DETAIL
ALL IMAGE DETAIL
USER FEEDBACK
USER SEARCH HISTORY
CATEGORIES
PEER
USERNAME
VIEWERS
EDIT
DELETE
NODE_A
pavithra