

# A Content-Aware Collaborative Accessing Model in Ad Hoc Networks

Bo Yang

Department of Computer Science  
Bowie State University  
Bowie, MD 20715 USA  
byang@bowiestate.edu

Manohar Mareboyana

Department of Computer Science  
Bowie State University  
Bowie, MD 20715 USA  
manohar@cs.bowiestate.edu

**Abstract**—Reducing the transmission cost is one of the most important issues in multi-hop ad-hoc networks. Considerable research has been done on efficient routing and data accessing algorithms. However, most previous research on routing schemes focus on topological structures, and not much attention has been made to data content distributions in ad-hoc networks. The exponential emergence of multimedia applications puts further challenges to the routing schemes. Due to the restrictions of ad-hoc networks and the volume of multimedia data, efficient data accessing schemes are needed for various applications. Within the scope of ad-hoc network paradigm, a semantic-based framework – extended summary schemas model (ESSM) – is presented and analyzed. Within the scope of this new framework several techniques are further proposed to optimize the cost of searching multimedia data in ad-hoc networks. The proposed model is simulated and the simulation results are analyzed.

**Keywords**- content-based retrieval, collaborative data accessing, ad hoc networks

## I. INTRODUCTION

Technical advances and new standards in wireless communications have stimulated the recent growth of wireless *ad-hoc* networks. An *ad-hoc* network is a collection of cooperative mobile nodes that communicate with each other without the intervention of infrastructures (centralized accessing points). Due to the lack of infrastructures, mobile nodes may not be directly connected to other nodes in the system. Hence, most research assumes *ad-hoc* network communications in a hop-by-hop fashion, where nodes behave as routers [2]. Stimulated by this unique communication pattern, various routing algorithms (proactive or reactive) are proposed to dynamically find reliable paths between mobile nodes [2]. However, research on routing schemes mainly focused on the topological structures, and not much attention has been made to data content distribution in *ad-hoc* networks. As an important hint of network status, data distribution may provide a novel guideline of routing and retrieving data in *ad-hoc* networks. In practical applications, several advantages can be achieved when considering the data distribution. We look at two different sample applications to motivate our research and the importance of data content analysis:

- 1) Distributed data sources: Consider the information sharing among the members of a police team in an anti-criminal action. Each member carries a mobile computer that is capable of receiving and sending images. At the same time, some of the team members may have the information of possible suspects – name, date of birth, height, blood type, and etc. Since the resources of mobile computers are limited, each computer may only contain a part of the complete list of criminals. If a suspect is found, more detailed information about this person may be needed for identification and proper action. Although, the picture of this person can be obtained immediately, the detailed information may not be readily accessible. And even worse, the location of the mobile host that contains the needed information could be an unknown question. Real-time determination of data source location is difficult in the traditional routing algorithms [11]. Without the knowledge of data distribution, broadcasting/multicasting seems to be the only choice. However, due to the sheer size of the multimedia data, broadcasting will drastically consume system resources – memory space, network bandwidth, and battery power. In contrast, large amount of resources can be saved if the data source is accurately located. The knowledge of data sources requires organizing *ad-hoc* network according to the data distribution.
- 2) Centralized data source: An *ad-hoc* network may comprise a commander and a group of soldiers. Each soldier carries a laptop, while the commander has a more powerful mobile computer connected to the satellite – the commander is clearly known as the data center. When a target is detected, the soldiers may need to request the commander for more detailed information, for the purpose of threat identification and further actions. However, this platform is not robust, scalable, or fault tolerant since the commander behaves as the data center and its movement within the area could also increase the network traffic. Operational support such as the *nearest neighbors* in response to a request reduces the network traffic and access latency; hence, a need for data-distribution-aware organization in the *ad-hoc* network.

---

This work was supported in part by the Chesapeake Information Based Aeronautics Consortium (CIBAC)

With the rapid growth of the wireless communication standards such as IEEE802.11, multimedia (image, audio, video, and etc.) applications are becoming more and more popular in mobile networks. Multimedia data are more powerful in expressing human thought than text information; however, they require large storage space, more computation capability, and higher network bandwidth. Indexing on multimedia data is rather difficult, which makes accessing or semantically organizing multimedia data more difficult to realize [3]. In *ad-hoc* networks, the autonomy and heterogeneity of mobile nodes introduce additional complexity to efficient representation and manipulation of multimedia data while minimizing computation and communication cost.

In this paper, we introduce a novel semantic-based multimedia-accessing scheme – Extended Summary Schemas Model (ESSM) – as the platform for multimedia retrieval. The hierarchical structure of the ESSM meta-data organizes multimedia data based on concise, abstract description of their semantic contents and allows content-based navigation through multimedia data sources. With the help of the ESSM meta-data, the multimedia retrieval in *ad-hoc* networks becomes a clearly aimed searching process that offers reduced network traffic, energy consumption, and response time, regardless of the distribution, heterogeneity, and autonomy of the multimedia data sources.

The remaining part of this paper is organized as follows: Section 2 overviews the background materials. Section 3 introduces the representation and manipulation of multimedia data within the scope of the summary schemas model. Section 4 addresses the clustering of multimedia objects in *ad-hoc* networks based on semantic contents. Section 5 analyzes the performance of the proposed platform. Finally, section 6 draws the paper to a conclusion.

## II. BACKGROUND

### 2.1 Content-based multimedia retrieval

Multimedia data sources usually comprise large volume of data whose contents cannot be defined and organized efficiently using traditional text-based indexing models [3]. The demand of multimedia data management, as a consequence, has motivated the research on content-based retrieval models [3]. The content-based multimedia retrieval models that have advanced in the literature can be categorized into three classes: partition-based models, region-based models, and human/computer integrated models.

The partition-based indexing models recursively divide the multimedia objects (or  $k$ -dimensional feature spaces) into disjoint partitions, with clustering or classifying algorithms, while generating a hierarchical indexing structure on these partitions. The earlier models in this class include Quad-tree [4], K-D-tree [5], and VP-tree [5]. The recent research has focused on models based on clusters [6]. The partition-based models normally employ very complex computations on features, which make these models inefficient for real-time multimedia applications. Moreover, the partition-based models do not consider feature spaces of different dimensions – i.e., multimedia objects from heterogeneous data sources [5].

The region-based indexing models employ small regions (either in the form of Minimum Bounding Rectangles or Spheres) to cover all the multimedia objects. Based on these bounding regions, a balanced tree is constructed as the indexing structure. This class of indexing models includes the R-tree and its variations (R\*-tree, R<sup>+</sup>-tree) [7], and SR-tree [8]. Relative to partition-based models, this indexing model improves storage utilization by avoiding forced splits. However, due to the shapes of the regions, this model has the born weakness of overlapping [7]. Furthermore, the data objects grouped in the same region may not share common semantic contents.

Some recently proposed models in the literature try to make use of human intelligence to facilitate accurate indexing. Benitez et al. [9] exploited the manual annotations to extract semantic contents. Cox et al. [10] used relevance feedback from users to improve the accuracy of similarity comparisons. Because of the human intervention in query processing, this class of indexing models achieves high accuracy in selecting similar multimedia objects. However, because of the very same reason, this approach requires comparatively longer response time and hence, is not suitable for real-time applications.

Due to the node mobility and data distribution, the traditional content-based retrieval methods cannot guarantee satisfactory performance in *ad-hoc* networks: First, unlike the wired systems, *ad-hoc* networks consist of mobile nodes that are not limited to certain spatial locations. Hence, it is not practical to partition the mobile nodes into spatial regions for purpose of efficient retrieval. Second, the traditional content-based retrieval models usually employ quite complex computations (such as minimum bounding rectangles, spheres, and vantage points), which puts unbearable burden to the mobile nodes. Finally, as shown in the aforementioned examples, data distribution plays an important role in multimedia retrieval in *ad-hoc* networks. The traditional models do not consider data distribution, which leads to the increased searching cost.

### 2.2 Object detection and recognition

Object detection and recognition have been playing important roles in multimedia applications such as video surveillance, object-based information retrieval, and human / computer interaction. The main challenge to object detectors is how to efficiently identify objects of various visual appearances from diversified backgrounds [15]. Many effective object-recognition algorithms and techniques have been reported in recent years. Considering the application domains and recognition rationales, the object-recognition methods as advanced in the literature can be divided into two main categories: invariant-shape recognition and flexible-shape recognition.

The invariant-shape object-recognition models focus on the detection and identification of objects of fixed shapes, such as face detection and digit recognition. Previous work on invariant-shape object recognition normally employs discriminative classifiers to facilitate fast detection. Scheinderman et al. adopted Adaboost confidence-weight prediction to minimize the classification error [15]. Viola et al.

used cascaded classifiers to reduce computation complexity by quickly discarding the background regions [16].

The need for detecting objects of various shapes and stances motivated the research on flexible-shape object recognition. Recent studies in the literature have proposed using either probabilistic models (e.g. Markov models and trees) or discriminative approaches (e.g. Support Vector Machine) to determine the contents of multimedia data. Ioffe et al. used mixtures of tree-structured probabilistic models for human tracking [18]. Felzenszwalb introduced MRF-based pictorial structures to represent the spring-like connections between object parts [17]. Mahamud et al. proposed combining simple discriminators for a multi-class classifier [19]. In contrast with invariant-shape object recognition, these models provide a way to general-purpose detection of objects from multimedia data collection at large.

In practice, the object-recognition methods made it possible to grasp the multimedia contents at an object level. Since one of the ultimate goals of object recognition is to provide content-based access of multimedia data with high accuracy and efficiency, the discussion in this paper will focus on representation and organization of multimedia data that make full use of the object-level contents obtained from object recognition.

### 2.3 Summary schemas model

The Summary Schemas Model (SSM) is an intelligent search engine that allows imprecise access to data sources in a multi-database environment. It is a global layer sitting on top of multiple preexisting autonomous and heterogeneous local databases [1].

The ESSM fulfills the indexing task through its hierarchical meta-data — summary-schemas hierarchy — which comprises three major components: a thesaurus, a collection of autonomous local nodes, and a set of summary-schemas nodes (Figure 1).

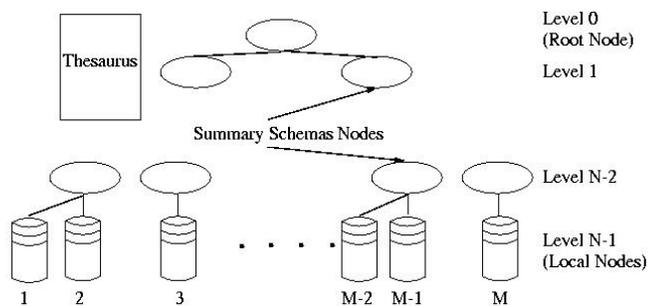


Figure 1. The summary schemas model.

SSM organizes the multimedia objects in a hierarchical fashion based on synonym, hypernym, and hyponym relationships among these objects. Synonyms are semantically similar objects in different data formats or same-format objects in different physical locations. A hypernym is the generalized description of the common characteristics of a group of multimedia objects. As the counter concept of hypernym, a

hyponym is the specialized description of the characteristics of multimedia objects. More details about ESSM can be found in [1].

### 2.4 Proposed methodology

Considering the aforementioned weaknesses of traditional content-based retrieval models, we propose a ESSM-based multimedia retrieval scheme for *ad-hoc* networks. The rationale of employing ESSM as the platform of multimedia retrieval is because of the following reasons:

- The ESSM provides comparatively more efficient content-based indexing capability: In the traditional content-based approaches, due to the lack of semantic analysis, semantically similar objects may be scattered in different regions, while semantically different objects may be grouped together. In contrast, ESSM represents a multimedia object according to its semantic content and clusters it with its semantically nearest neighbors, which drastically reduces the searching cost.
- SSM integrates heterogeneous data sources into a unified logical system: ESSM organizes multimedia objects regardless of their representation (data formats such as JPEG, TIFF, or MPEG); uniformly, according to their contents. In addition, different media types (video, audio, image, and text) can be integrated under the ESSM umbrella, regardless of their physical differences. Hence, ESSM provides good scalability and robustness in heterogeneous distributed environments such as *ad-hoc* networks.

In this paper, we present a multimedia retrieval mechanism for *ad-hoc* networks. The novelty of our approach stems from the following factors:

- Content-based organization of the ESSM allows users to submit imprecise queries. We have developed a Graphical User Interface (GUI) that facilitates submission of the queries.
- The query processing is restricted within a very small subset of all mobile nodes in an *ad-hoc* network. When a query is submitted, instead of traversing throughout the *ad-hoc* network, our approach selects a small subset of mobile nodes (a cluster): this cluster includes all and only the nodes that contain multimedia objects related to the query. The nearest neighbors are generated within this cluster, which drastically reduces the searching cost.
- A dynamically self-adjusting indexing structure — ESSM hierarchy — behaves as the spine for query processing. By summarizing the data contents of mobile nodes and organizing them into clusters, the ESSM hierarchy provides an effective way that dynamically identifies a cluster of mobile nodes for a query, regardless of their spatial locations in the *ad-hoc* network.

## III. SEMANTIC-BASED REPRESENTATION

In multimedia retrieval systems, there are two types of features: granule-level features and object-level features. The

granule-level features are those characteristics that directly or indirectly are derived from the original format of multimedia storage — i.e., the pixels, such as hue, textures, and saturation. The object-level features, in contrast, are obtained from the recognition of the higher-level understanding of the multimedia data — the semantic topics of the multimedia data. In this work, we analyze the semantic contents of multimedia objects by considering both types of features.

To represent the contents of multimedia objects in a computer-friendly structural fashion, the ESSM organizes the multimedia objects into a hierarchical organization according to their semantic contents. A multimedia object, say, an image, can be considered as the combination of a series of elementary entities, such as animals, vehicles, and buildings.

### Definition 1: The Elementary Entity

The elementary entities are those data entities that semantically represent basic objects (objects that cannot be divided further). Formally, the summary schema of an elementary entity ( $E$ ) can be considered as a first-order logic expression.

Let  $E = f_1 \wedge f_2 \wedge \dots \wedge f_n$ , where  $f_i = p_{i1} \vee p_{i2} \vee \dots \vee p_{im}$  is the disjunction of some logic predicates (true/false values) and  $p_{i1} \dots p_{im}$  form a logic predicate set  $F_i$ . — In the feature-based multimedia data sets,  $f_i$  indicates the  $i^{\text{th}}$  feature of the elementary entity. The summary schema of an elementary entity can then be defined as:

$$E = \bigwedge_{i=1}^n \left( \bigvee_{j=1}^m p_{ij} \right), \quad \text{for every } p_{ij} \in F_i.$$

Note that in any term  $f_i = p_{i1} \vee p_{i2} \vee \dots \vee p_{im}$ , there is one and only one true predicate  $p_{ij}$ . For instance, if  $p_{i1}, p_{i2} \dots p_{im}$  correspond to all possible color patterns, the semantic content of  $f_i$  at any time is a specific color pattern. Since  $f_i$  is disjunction of  $p_{i1}, p_{i2} \dots p_{im}$ , the false predicates do not affect the final result. The content of an elementary entity is restricted by its conjunction terms  $f_1, f_2 \dots f_n$ , which are the extracted features in application domains.

### Definition 2: The Multimedia Object

A multimedia object is the combination of a series of elementary entities. Given a multimedia object including  $k$  elementary entities  $E_1, E_2, \dots, E_k$ , the summary schema of this multimedia object can be defined as:

$$S = \text{opt} \left( \bigcup_{i=1}^k E_i \right),$$

where  $\text{opt}$  is a function that converts a logic expression into a semantically equivalent and shorter form.

The summary schemas provide a means to automatically define accurate description of complex objects: For a given multimedia object, say an image, a series of elementary entities can be recognized by utilizing some pattern recognition techniques [6]. These entities are then described using logic expressions, as noted in definition 1. Finally, the object is represented as the union of its contents (see Figure 2).

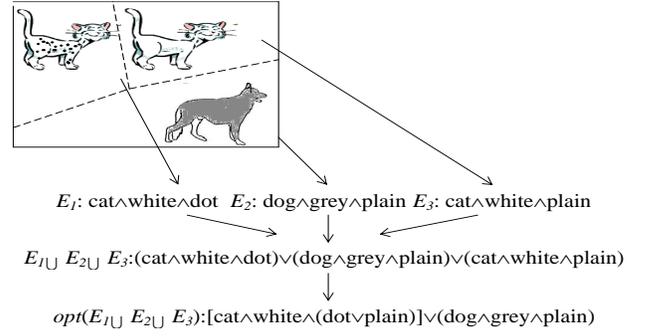


Figure 2. Summary schema of an image.

## IV. COLLABORATIVE CLUSTERING

The semantic-based representation presented in section 3 provides a paradigm to describe multimedia data. Based on this paradigm, contents of the multimedia objects on mobile nodes can be automatically identified, summarized, and expressed as logic equations with strong mathematical backbone (summary schemas). Logic representation of semantic contents of multimedia data sources would also allow simple and efficient recognition of similar entities which assists classification and clustering of multimedia data.

### Definition 3: Content-Related Nodes

Suppose a mobile node  $n_i$  contains multimedia data objects  $d_1, d_2, \dots, d_m$ , which are collectively denoted as  $D(n_i)$ . Based on the aforementioned ESSM representation of multimedia data object, each mobile node  $n_i$  can obtain a summary schema  $S(n_i)$  that abstracts the contents of data objects in  $D(n_i)$ . Given a pair of nodes  $n_0$  and  $n_1$ , they are content-related iff:

$$S(n_0) \oplus S(n_1) \neq S(n_0) \vee S(n_1),$$

where  $S(n_0) \oplus S(n_1)$  is defined as  $(S(n_0) \wedge \sim S(n_1)) \vee (\sim S(n_0) \wedge S(n_1))$ .

In another words, if nodes  $n_0$  and  $n_1$  are content-related, they must have some common data entities  $D(n_0) \cap D(n_1) \neq \emptyset$ .

### Definition 4: The Content-Based Cluster

Suppose an *ad-hoc* network  $N$  comprises  $k$  mobile nodes  $n_1, n_2, \dots, n_k$ . Let  $n_i \approx n_j$  denote the content-related relationship between  $n_i$  and  $n_j$ , and  $n_i \not\approx n_j$  denote that  $n_i$  and  $n_j$  are not content-related. Then a content-based cluster  $C$  is defined as follows:

$$C = \{ n_i \mid \forall n_j \in C, n_j \approx n_i; \forall n_k \notin C, n_k \not\approx n_i \}$$

Based on the summary schemas generated from each node, the *ad-hoc* network is partitioned into clusters where each cluster contains mobile nodes with similar or overlapping data objects. The summary schemas within each cluster are integrated to form a summary schema that represents the semantic contents of each cluster. Cluster level summary schemas are then integrated and fused together (based on their semantic similarity) to form higher level clustering. This process is recursively performed until the whole network is represented as one cluster.

Figure 3 represents a collection of mobile multimedia data objects located irregularly on the mobile nodes in the *ad-hoc* network. As can be observed, the semantically similar contents (say, cars) are scattered on nodes not spatially located together. However, these nodes can be treated collectively as a cluster based on the aforementioned discussion. For instance, node A, B, D, and J can be clustered together because of their semantic contents. Figure 4 depicts the application of our clustering procedure and formation of the summary schemas hierarchy on Figure 3. Note that in this organization, content-based clusters are semi-isolated from each other; as a result, queries based on their semantic contents are directed to the most relevant clusters rather being flooded through out the *ad-hoc* network.

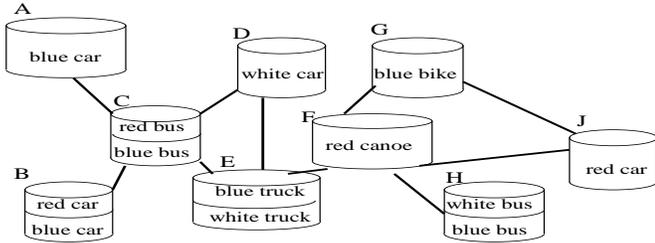


Figure 3. Data distribution in an *ad-hoc* network.

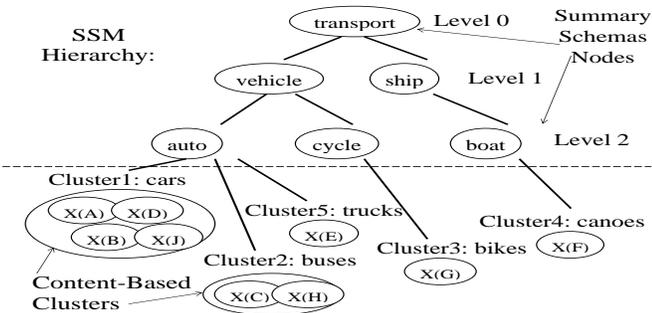


Figure 4. The ESSM hierarchy based on clusters.

## V. PERFORMANCE EVALUATION

### 5.1 Theoretical analysis

The performance of the proposed scheme is impacted by several factors as such; the initialization overhead of building the indexing hierarchy (including clustering and centroid selection), the cost of performing content-based multimedia queries, the overhead of maintaining the indexing hierarchy when network status changes. To facilitate content-based retrieval, these factors need to be analyzed in details.

Our analysis is based on the following notations:

- $r$ : the number of nodes in the ad hoc network.
- $m$ : the minimum number of children for an indexing node (minimum fan-out).
- $P_j$ : the probability of a node joining the network.
- $P_R$ : the probability of a node leaving the network.
- $P_M$ : the probability of a modify operation.
- $Q$ : the query rate.

### A. Initialization Overhead

In the initialization step, a node with highest computational power and largest communication capability, say  $n_i$ , is chosen as the coordinator to construct the indexing hierarchy. The selection of the coordinator node takes  $\theta(r)$  hops. The coordinator needs to send each node a message to notify the coordinator's location and to collect data content descriptions, which takes  $O(r).r$  hops. In addition, the construction of the hierarchical infrastructure takes  $O(r)$  hops because the message cost is proportional to the number of edges (or links): the hierarchy has less depth than the binary search tree due to its minimum  $m$ , thereby possessing edges less than  $r+r/2+r/4+\dots = 2r-1$ . Hence, the initialization overhead amortized on each node is  $\frac{\theta(r) + O(r)r + O(r)}{r} = O(r)$  hops, where  $r$  is the number of nodes in the ad hoc network.

### B. Search Cost

Resolution of a query requires at most  $2\log_m(r)$  logical steps to traverse the indexing hierarchy. Each forwarding operation takes at most  $r$  hops. Hence the average searching cost for a query is  $O(r \log_m(r))$  hops. In contrast, the flooding strategy requires  $\Omega(r^2)$  hops to resolve a query, since a network of  $k$  mobile nodes can have  $\theta(r^2)$  connections, and a query may be transmitted on each connection at least once.

### C. Maintenance Overhead

As noted before, the proposed indexing hierarchy does not change as long as the data source contents are intact. As a result, the indexing hierarchy changes if a data source is inserted/deleted in/from the network (including the removal of the centroid node) or a modification is made to the contents of an existing node. As noted before, content changes on each node is piggybacked with unresolved query and forwarded to the cluster centroid to potentially trigger maintenance overhead due to the reconstruction of the indexing hierarchy. Since the whole network is connected through the indexing hierarchy, the deletion or removal of any centroid node can be detected when a query cannot be resolved locally and forwarded to the centroid node. Therefore, the query resolution also provides a mechanism for checking the connectivity.

A new node  $n_{r+1}$  joining the network requires at most  $r \log_m(r)$  hops to be included in its related cluster: joining any cluster from the bottom level of the hierarchy takes  $\log_m(r)$  hops, therefore comparing all possible clusters takes  $r \log_m(r)$  hops. Hence, the average cost for processing new nodes is  $(P_j) Q r \log_m(r)$  hops: with query rate  $Q$ , the node joining event happens  $(P_j) Q$  time, each time takes  $r \log_m(r)$  hops. Similarly, the processing of leaving nodes requires  $(P_R) Q r \log_m(r)$  hops. The modification operation can be viewed as a deletion of a node followed by an insertion of a new node. Consequently, it requires  $((P_M) Q r \log_m(r) + (P_M) Q r \log_m(r))$  hops. Finally, the average modification cost is  $\theta((2P_M + P_j + P_R) Q r \log_m(r))$ .

### 5.2 Experimental evaluation

In this section, we present the experimental analysis of the proposed content-based clustering model (called Extended SSM or **ESSM** in the following discussion) against the

flooding-based search scheme. As noted earlier, in an ad hoc network with distributed multimedia data sources, flooding-based blind search strategy may cause extra computational and communication overheads due to the forwarding of the queries to every node. In contrast with the flooding-based schemes, the ESSM organizes nodes with similar contents into clusters, and forwards query packets only to the content-related nodes.

The evaluation consists of a series of experiments conducted using both real data set and synthetic data set. Our comparative analysis is based on various performance metrics such as accuracy, search cost, scalability, and physical characteristics of the mobile nodes.

#### A. Experimental Setup

The experiments were run on the basis of ad hoc network prototype with CMU extension to the ns-2 version 2.26. Since ns-2 does not support content-based information retrieval, a semantic-representation module was developed and added to facilitate multimedia data organization. In addition, the routing and data transmission processes were implemented under the framework of extended summary-schemas hierarchy.

The experiment was initialized by assuming a default number of pre-existing nodes in the network with random connectivity among the nodes. A mixture of operations, including querying, node joining, and node leaving, were randomly generated and submitted to the network. The query generation time follows the exponential distribution, which is similar to the previous work [11]. The access pattern in the queries follows *Zipf-like* distribution, which is widely used to model non-uniformly distributed queries [12]. The input parameters are summarized in Table 1. Most of the parameters are self-explanatory. More details for some parameters are given below.

**Node Movement Parameters:** Each node in the experimental environment randomly selects waypoints within a 670m \* 670m flat area. The node density can be adjusted by changing the number of mobile nodes in the range of 1 to 16,384 in the flat area. The node movement pattern follows the random waypoint movement model: Initially, the nodes are placed randomly in the area. Each node selects a random destination and moves toward the destination with a speed selected randomly from  $[0, v_{max}]$ . After reaching its destination, the node pauses for a period of time and repeats this movement pattern.

**Dataset Parameters:** To examine both the accuracy and the scalability, we used two sets of experimental datasets as the test beds — the synthetic dataset and the real dataset as follows:

- The synthetic dataset employed in the experiments is similar to the one used in [1], which includes up to 65,536 data points in a 256-dimension feature space whose feature values are assigned by a random number generator abiding normal distribution in the interval  $[0, 1]$  on each dimension.
- The real dataset we used consists of 2,560 images of 32 semantic categories from the Corel dataset. Each image in the dataset is represented as a vector of 256 features (color histograms and texture wavelet coefficients) and 4 annotation keywords. It is a large and heterogeneous image dataset. 2,048 images in the test dataset were used to train a

semantic subspace learning module (i.e. LPP) integrated in the ESSM system that deduces the relationships among the keywords and the semantic categories. The extra 512 images were used as the candidate pool for query examples.

TABLE I. INPUT PARAMETERS TO THE EXPERIMENTAL SYSTEM

<i>Parameter</i>	<i>Value</i>
Number of nodes	128
Image dataset size	2,560
Texture feature type-1	Daubechies wavelet coefficients
Texture feature type-2	Tamura coarseness histogram
Texture feature type-3	Pyramid wavelet coefficients
Color feature type-1	Color histogram in HSV space
Color feature type-2	Color coherence vector in LUV
Color feature type-3	First and second moment in Lab
Environment size	670m * 670m
Transmitter range	100m
Node mobility ( $v_{max}$ )	1 m/s
Pause time	1s
Bandwidth	1M bps
Query packet size	512 bytes
Query rate ( $Q_{rate}$ )	10 query/s

#### B. Retrieval Performance Evaluation

As noted earlier, the simulator is intended to evaluate the performance improvement of the proposed clustering scheme based on metrics such as search cost and accuracy:

- **Accuracy** is the percentage of the results generated by the search schemes (either the proposed clustering model or flooding) matching the results of centralized strategy. The higher matching percentage implies higher retrieval accuracy.
- **Search path length** is the average number of hops spent on locating the data source node that contains the semantically nearest neighbors of the query image.
- **Search cost** is the average number of messages incurred during the process of query resolution. Flooding-based schemes may have short search path length, but their search cost is high due to the duplicate query replies.
- **Maintenance cost** is the number of messages spent on adjusting the indexing hierarchy according to the topology and content changes in the network.

#### Retrieval Accuracy

The retrieval accuracy is evaluated for different simulation settings, varying the number of mobile nodes, the speed of node movement, the rate of query submissions, and the number of hops, during the query resolution process. In these experimental runs, we assumed 2,048 real images randomly scattered on mobile nodes. The maximum fan-out of the indexing hierarchy was set to 10, and the number of nearest neighbors in the content-based search was set to 10.

Figure 5 illustrates the impact of the number of hops on the retrieval accuracy for a network of 128 mobile nodes moving at maximum speed of 1m/s. Given a fixed setting of network scale (128 mobile nodes), the number of hops was varied from 10 to 80. As can be observed the retrieval results generated by the ESSM are more semantically related with the query image. Also note that both schemes (ESSM and flooding) achieves better accuracy as queries are generated at lower rate. This can be explained by the following fact: as query rate increases, more query packets are added into the network traffic and thereby causing more packet loss, and more hops are required for resending and forwarding the queries. Therefore, the query results obtained from large hop counts are removed from the result list due to the limitation of hops in Figure 5. The superior performance of the ESSM in contrast with flooding stems from its content-based clustering capability, since the search domain is restricted to a few clusters that are semantically most relevant to the query.

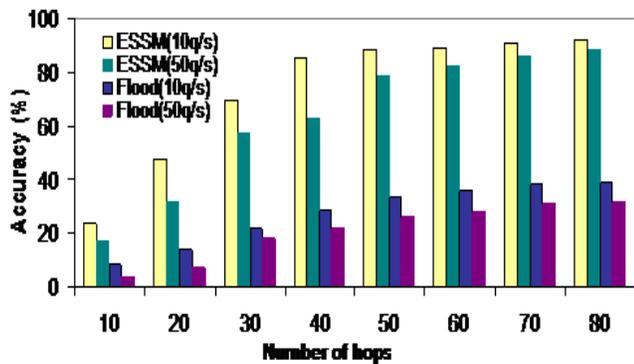


Figure 5. Impact of hop count (k = 10).

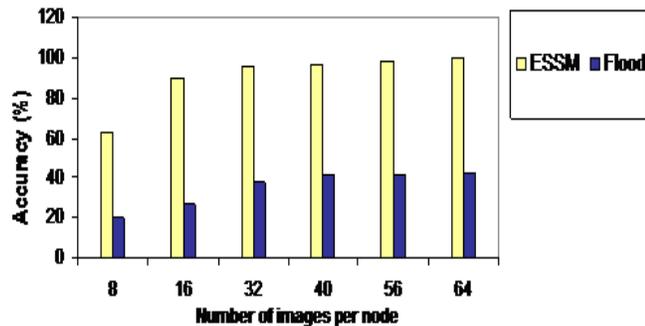


Figure 6. Impact of data density (query rate = 10q/s, k = 10).

In a similar simulation run, we examined the impact of data density on the search accuracy. Using the same real image dataset, we varied the number of mobile nodes from 256 down to 32, increasing the average number of images per node from 8 to 64. With the fixed mobility and query rate (moving speed is limited to 1m/s and an average of 10 queries per second), the number of hops was limited to 64. Figure 6 depicts the experimental results. Since the probability of finding semantically related results increases as the data density increases, both schemes offer better accuracy as data density

increases. However, the accuracy of flooding-based schemes is still less than that of the ESSM. Also note that the ESSM achieves over 90% search accuracy at a relatively low data density (16 images per node). This implies that in a relatively large-scale network (2048/16 = 128 nodes), the ESSM almost achieves its peak performance with a small search cost (less than 64 hops). In the scope of wireless ad hoc networks, this is a significant observation since the mobile nodes usually have small storage and cannot support large data density.

### Search Cost

In an ad hoc network, the search path length is usually calculated as the number of hops needed to deliver the query to proper mobile nodes that contain requested data. The real image dataset was used in this simulation run. Several factors, such as mobility, node density, and query rate, have direct impacts on the search cost; hence, we ran our simulator with different combinations of these parameters.

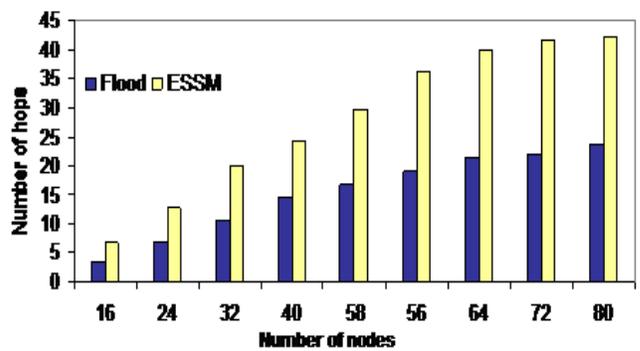


Figure 7. Comparison on search path length (query rate = 10q/s, k = 10).

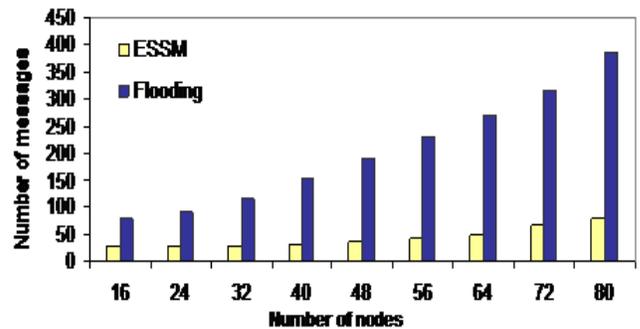


Figure 8. Comparison on search cost (query rate = 10q/s, k = 10).

Figure 7 illustrates the number of hops needed during the query resolution for various network densities. As can be seen from Figure 7, the flooding scheme finds the data source node using less number of hops. This is due to the fact that flooding broadcasts the query to all its neighboring nodes and reaches the data source node using the shortest path. However, due to the broadcasting and duplicate replies, the search cost of flooding may be formidably high (see Figure 8).

Figure 8 shows the search cost of both schemes in the same environment. As anticipated the ESSM resolves a query much

faster than flooding-based scheme due to its content-based clustering characteristic. In addition, from Figure 8 it can also be concluded that as node density increase, the ESSM offers even better performance than the traditional flooding schemes — steady performance improvement as the network scales up. The better performance of ESSM in comparison to flooding can be explained by the fact that larger node density will incur heavier flooding overhead, while ESSM can utilize the denser node distribution to form larger size clusters, which increases the probability of resolving a query within one cluster.

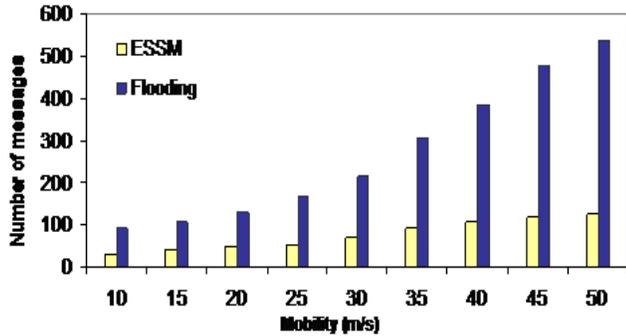


Figure 9. Impact of mobility (query rate = 10q/s, k = 10).

We also evaluated the impact of node mobility on search cost. For an ad hoc network of 64 nodes the maximum speed of mobile nodes varied from 10 to 50 m/s. Figure 9 shows the simulation results. As can be concluded, the search cost of both schemes increases as the node mobility increases, since increased mobility can cause more frequent breaks of the connectivity among mobile nodes, leading to higher cost in creating and maintaining the routes; however, the ESSM resolves queries at comparatively much less cost than flooding. This is due to the very nature of the flooding scheme — higher network traffic and higher workload at mobile nodes. In contrast, the ESSM resolves the queries within the scope of content-related nodes, and avoids forwarding queries to irrelevant nodes.

### Scalability

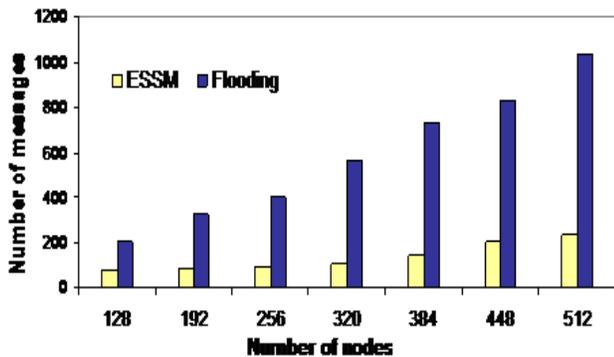


Figure 10. Search cost in large-scale networks. (query rate = 10q/s, k = 10)

In a separate simulation run, the search cost of both schemes was evaluated as the network scales up. In this simulation run, we varied the number of nodes from 128 to 512 and randomly distributed 65,536 synthetic data points on the mobile nodes. Figure 10 illustrates the result. Similar to our earlier observation (Figure 8) one can conclude that the ESSM is scalable to large network sizes and large number of data objects.

### Maintenance Overhead

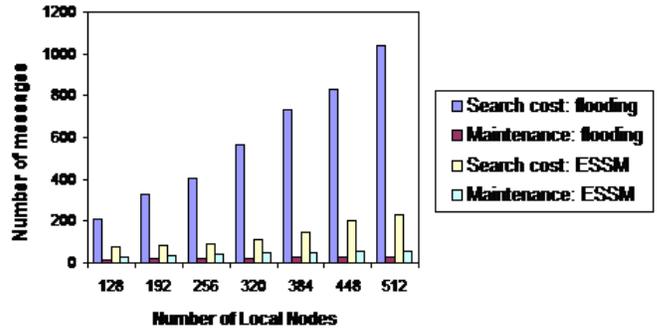


Figure 11. Search cost and maintenance cost (65,536 data points, mobility = 1m/s, query rate = 10q/s, k = 10).

As mentioned before, self-adjustment capability of the proposed model incurs maintenance overhead that needs to be evaluated. In contrast to the ESSM, the maintenance cost of flooding scheme is very limited to the messages needed to update the neighborhood relationships between mobile nodes. Figure 11 shows the average search cost and the maintenance cost of both strategies as the network becomes denser. As one can conclude, even taking the maintenance overhead into account, the ESSM still offers a better overall performance than flooding.

## VI. CONCLUSIONS

Data retrieval in ad-hoc networks is a challenging task due to the node mobility and the lack of infrastructure. The study of content-based multimedia retrieval puts further challenges to the research of data management in *ad-hoc* networks. In this paper we presented a semantic-based multimedia accessing model – ESSM – to reduce the searching cost in *ad-hoc* networks. It first clusters mobile nodes based on their data contents, then summarizes the data within a cluster, and finally forms a hierarchy of summary schema nodes to organize the data within the whole network. By carefully mapping summary schemas nodes to proper mobile nodes, this model avoids adding extra physical nodes or increasing the overhead of network traffic.

With the summary schema hierarchy, the content-based multimedia retrieval in *ad-hoc* network becomes a fully aimed and determinate process that is performed within a small range of nodes. The query processing is performed by forwarding request packets only to the data sources most relevant to the queries, instead of flooding the network. Our simulation shows that ESSM approach achieves better performance than

broadcasting method based on performance metrics, such as mobility, number of nodes, and cache size.

#### REFERENCES

- [1] M. W. Bright, A. R. Hurson, and S. Pakzad. Automated Resolution of semantic heterogeneity in multidatabases. *ACM Transactions on Database Systems*, 1994, 19(2): 212-253.
- [2] J. Broch, D. Maltz, D. Johnson, Y. Hu, and J. Jetcheva. A performance comparison of multi-hop wireless ad hoc network routing protocols. *ACM Mobicom*, 1998: 85-97.
- [3] G. Auffret, J. Foote, C. Li, B. Shahraray, T. Syeda-Mahmood, and H. Zhang. Multimedia access and retrieval: The state of the art and future directions, *Conference on multimedia*, 1999, pp: 443-445.
- [4] H. Samet. The Quadtree and related hierarchical data structures. *ACM computing surveys*, 1984, 16(2): 187-260.
- [5] A. W. Fu, P. M. Chan, Y. Cheung, and Y. S. Moon. Dynamic VP-tree indexing for n-nearest neighbor search given pair-wise distances. *VLDB Journal*, 2000, 9(2): 154-173.
- [6] J. B. Kim and H. J. Kim. Unsupervised moving object segmentation and recognition using clustering and a neural network. *Proceedings of the 2002 International Joint Conference on neural networks (IJCNN '02)*, 2002, 2: 1240-1245.
- [7] N. Beckmann, H. Kriegel, R. Schneider, and B. Seeger. The R\*-tree: An efficient and robust access method for points and rectangles, *Conference on Management of data*, 1990, 19(2): 322-331.
- [8] N. Katayama, S. Satoh. The SR-tree: An index structure for high-dimensional nearest neighbor queries, *conference on Management of data*, 1997, 26(2): 369-380.
- [9] A. B. Benitez. Semantic knowledge construction from annotated image collections. *Multimedia and Expo*, 2002, 2: 205-208
- [10] I. J. Cox, M. L. Miller, T. P. Minka, T. V. Pappathomas, and P. N. Yianilos. The Bayesian image retrieval system, PicHunter: theory, implementation, and psychophysical experiments, *IEEE Transactions on Image Processing*, 2000, 9(1): 20-37.
- [11] C. E. Perkins, E. M. Royer, S. R. Das, M. K. Marina. Performance comparison of two on-demand routing protocols for ad hoc networks. *IEEE Personal Communications*, 2001, 8(1): 16-28.
- [12] M. Thottan, V. Mirrokni, H. Uzunalioglu, and S. Paul. A simple polynomial time framework for reduced path decomposition in multi-path routing. *IEEE Infocom*, 2004.
- [13] E. M. Royer and C. E. Perkins. An implementation study of the AODV routing protocol. *IEEE Wireless Communications and Networking Conference*, 2000.
- [14] T. Guven, C. Kommareddy, R. La, M. Shayman, and S. Measurement based Optimal multi-path routing, *IEEE Infocom*, 2004.
- [15] H. Schneiderman and T. Kanade. Object detection using the statistics of parts. *International Journal of Computer Vision*, 2002.
- [16] P. Viola and M. Jones. Robust real-time object detection. *International Journal of Computer Vision*, 2001.
- [17] P. F. Felzenszwalb and D. P. Huttenlocher. Efficient matching of pictorial structure. *IEEE Conference on Computer Vision and Pattern Recognition*, 2000: 66-73.
- [18] S. Ioffe and D. Forsyth. Human tracking with mixture of trees. *IEEE International Conference on Computer Vision*, 2001(1): 690-697.
- [19] S. Mahamud, M. Hebert, and J. Lafferty. Combining Simple Discriminators for Object Discrimination. *European Conference on Computer Vision*, 2002.