ISSN 2047-3338

A Survey on Caching, Replication and Crowdsourcing in Mobile Environment

Kshama Raichura[1], Nilesh Padhariya[2] and Kishor Atkotiya[3]

***Abstract*—In this new era of technology, the mobile devices are available with various features. The main functionality of the mobile devices became effective *information sharing* i.e. information access on every finger touch, which makes the whole world small. The information sharing is done through various platforms like social networking, cloud computing, peer-to-peer communications and so on. Nowadays the mobile phones are not used just for basic communication as they are available with various features same like any other computing devices. For example, a mobile phone has capacity to store various databases and facility to access information over the networks. The main contribution of this survey include (a) the applicability of mobile databases, (b) the caching and the replication in mobile environment, and (c) the crowdsourcing i.e. outsourcing in the crowd.**

***Index Terms*—Caching, Replication, Crowdsourcing, Mobile Databases and Mobile Environment**

I. Introduction

TODAY'S world revolutionized the communication devices from the bigger telephonic medium to the tiny but powerful computing and information sharing devices. They are handy and so easily movable; hence those devices are called mobile devices. Furthermore, the increased usage of mobile devices causes the exponential increase in information network size, which effectively becomes a mobile environment.

Information Telecommunication Union (ITU) [1] gathers data about the mobile subscriptions registered across the globe and generates the various statistical reports, which show the proliferation of mobile devices. Based on ITU's reports, Figure 1 depicts the growth of mobile cellular subscriptions over past 12 years (2000-2011) in the developed and developing countries. In year 2000, mobile usage in the developing countries was almost half of that in the developed countries. Interestingly, by 2005, developing countries' mobile usage just crossed the developed countries' mobile usage and by 2011, this ratio became just inverse by 3 times i.e., the mobile usage in the developing countries is just thrice of the mobile usage in the developed countries. According to the telecommunication data almost 87% of population has mobile devices in today's era.

Figure 2 represents the worldwide growth of mobile cellular subscriptions per 100 inhabitants. In 2001, the growth was comparable in the developed and the developing nations. This was because the cellular infrastructure was not much

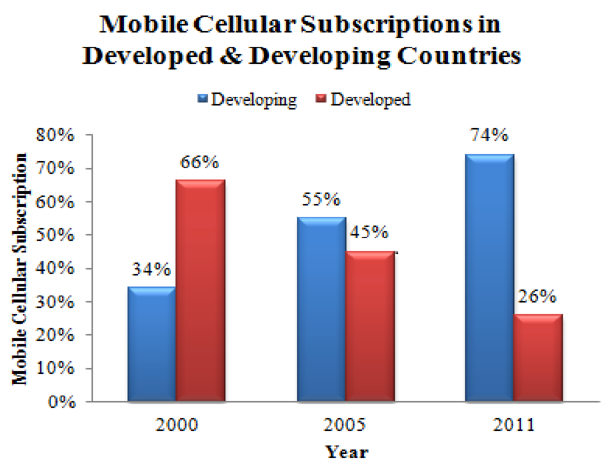

Fig. 1. Mobile cellular subscriptions over the globe

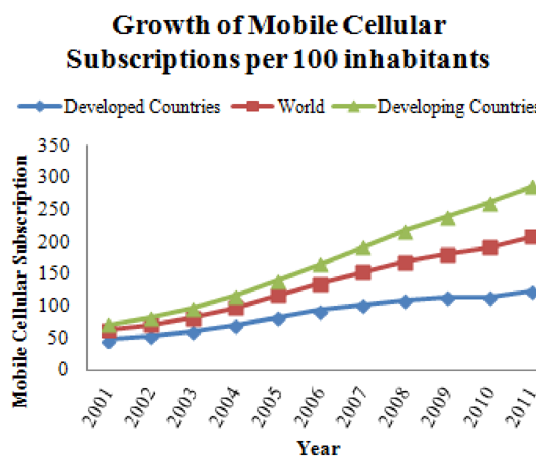

Fig. 2. Growth of mobile cellular subscriptions

[1]Shree M. & N. Virnai Science College, Gujarat, India, (Email: kdraichura@vsc.edu.in)
[2]Atmiya Institute of Technology & Science, Gujarat, India, (Email: nilesh@aits.edu.in)
[3]Atkotiya, Saurashtra University, Gujarat, India, (Email: atkishor@yahoo.co.in)

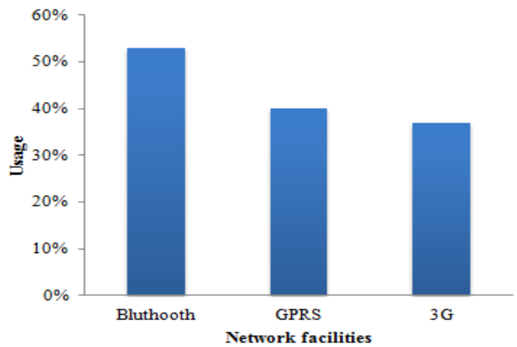

Fig. 3. Usage of different facilities in mobile devices

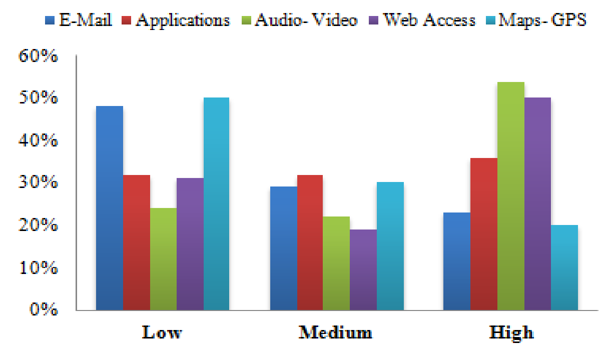

Fig. 4. Priorities of various applications available in mobile devices

established in the developing nations as compared to that in the developed nations. As the communication infrastructure became wider and stronger; and the availability of mobile devices became higher, the growth in the developing nations became almost double than that in the developed nations.

To explore the recent mobile usage across the society, we have carried out a small-scale survey, which shows the significant usage of mobile systems in various fields such as social networking, information sharing, entertainment etc. This survey was performed using the cross sectional method. Here, we present the related results and the analysis of those results. We have collected data of around more than 6000 mobile users through online feedback system. Figure 3 depicts the results of the percentage of mobile users based on network facilities. Observe that, the large percentage of communication is done over the short-range technology Bluetooth, as it is more cost-effective and handy to use. Interestingly, the result also demonstrates that recently the mobile users are more relied on 3G network communications over GPRS due to their faster and more reliable connectivity.

Figure 4 presents the application-based mobile usage over three different priorities i.e., low, medium and high of the mobile users. Here, priority is based on the user's preference of applications. For example, the percentage of mobile users, who has high priority to E-mail system, is 22%, while the percentage of mobile users with low priority to E-mail system is 48%. The results show that mobile users have high priority for the applications related to audio-video and web-access compared to others. While low-priority applications include those applications, which rely on GPS like technology due to its more power consumption.

Moreover, increasing proliferation of mobile devices is there because of increase in availability of various facilities in recent mobile devices. Such provisioning on tiny device facilitates the ease of data transfer, effective gathering of geographical information (on the move), improving reliability of data through uploading and downloading on remote server, exchange of various types of data like text, image, audio, video.

The reminder of this paper is organized as follows. Section 2 presents the various application scenarios related to mobile environment. Section 3 discusses the various caching techniques for obtaining faster query responses in mobile environment. While, for improving data availability in mobile networks, the various replication schemes have been proposed, which are discussed in Section 4. Section 5 explains the scenario of message communication in mobile networks. Section 6 discusses the novice data management through crowdsourcing i.e., outsourcing in the crowd, using mobile technology. Finally, we conclude in Section 7 with directions for future work.

II. Mobile Applications

This section motivates the mobile computing by means of exploring several areas where the mobile devices become handy computing tools for effective analysis and rapid sharing of the information on-the-go. Here, we have focused on mobile applications related to information sharing, m-commerce, medical healthcare, Near Field Communication (NFC) etc.

A. *Information Sharing*

The main inspiration behind the rapid development of the internet application is information sharing. In this new era of technology, gradually our communication is getting reshaped because every finger touch gets information from the world due to internet availability. Through the information sharing, we daily get various information on our mobile devices like railway ticket reselling, carpooling information and many short term discount schemes. Such type of information may be available on our mobile devices but not spontaneously. Hence, it may be possible that information is outdated and may not be useful.

The work in [2] demonstrates Thumb, a system to share information instantly on smart phones, especially for resources with extreme short life time. Using Thumb, a third-party can share the information with others, while the users get spontaneously notification about that. For example: any passenger wants to cancel the railway ticket just before one day of the train departure, and wants to sale ticket, then that user is able to advertise this information over the internet.

B. *Near Field Communication (NFC)*

Due to nonstop development of technology, the digital world becomes border less. Nowadays, e-Transaction widely adopted to simplify the process of transaction across various platforms such as e-ticket, e-commerce, etc. Moreover, the virtual money is involved in every virtual transaction. For example, shopping with virtual money becomes a common practice using NFC, especially, when we have smart phones.

A novice prototype for Train Ticket Application using NFC is shown in [3]. In this system, with the help of NFC enabled mobile device, the passenger gets the information about the vacant seats in the train and passenger feels like he/she purchasing this ticket because, a passenger can get the scanned copy of the e-ticket on the mobile device, when he/she completes the payment using the voucher system. When conductor approaches the passenger in the train, the passenger sends ticket data to the conductor and once the conductor receives the data from passenger, the ticket data will be destroyed from the passengers' device.

C. Healthcare

In developing countries, the rural areas still may not benefited by the basic civilized facilities, especially, the health care and medications. Due to inappropriate healthcare facilities, sometimes people may lose their life in the rural areas. Wearable sensors are the solution for this. The wearable sensors can continuously monitors the patients and issues the warnings to the doctors or care takers by sending messages on their mobile devices.

The content in [4] represents the decision support system which collects data from various wearable sensors and analyzed this data for the variety of diseases. Furthermore, this result will be stored and sent to the required person as a Short Message Service (SMS).

The proposed system in [5] integrates the wearable sensors with mobile device and developed a platform which continuously monitors the patient. In case of emergency, this system is capable of sending SMS on the doctor's mobile phone. This system collects data from the various wearable sensors in the mobile, analyze that data and send it to the centralized server.

D. Mobile Databases

Mobile database system is defined as a distributed database system, where the database may be placed or scattered across the mobile hosts i.e., capable to communicate via short-range technologies such as WiFi or Bluetooth [6]. Due to proliferation, preferential and propagation of mobile marketing, information flow need to be managed and directed intelligently for the support applications for effective data management. In mobile environment, data is distributed across the network majorly in the form of pushing (data dissemination) and pulling (data hosting) techniques. Few of the possible techniques based on push/pull mechanism have been surveyed in [7]. This includes data dissemination over limited bandwidth channels, location-dependent data querying and advanced interfaces for mobile computers. Moreover, these approaches have few issues to be addressed such as how to store information on database, how to effectively and optimally design database for mobile network, low bandwidth in network; and frequent network partitioning caused by mobility and energy constrained of peers. The works in [8], [9], [10] show the several approaches to address these research issues in mobile networks.

Furthermore, the work in [11] presents the new execution scheme and concurrency control mechanisms to take full advantage of mobile databases, where the transaction management deals with number of their related constraints. Similarly, in mobile environment, each mobile host has limited local storage capacity, their mobility constraints to distribute the copy of datasets among mobile hosts, decentralized causes hard to locate and access data. [8] also proposed an approach to effectively utilize the limited mobile storage to store mobile databases using compression technique that ignores the semantics of data, known as Database Summarization.

Moreover, [11] reports study on multi-version transaction processing approach and deadlock-free concurrency control mechanism based on multi-version two-phase locking scheme integrated with timestamps approach. They also performed the comparison of schemes using reference model by observing the behaviour of the proposed model with simulation study in a mobile environment. The outcome of the experiments shows that this model reduces the response time and minimizes the restarts and aborts. Thus, it improves the degree of concurrency and provides significantly higher throughput. [8] also proposed an approach to effectively utilize the limited mobile storage to store mobile databases using compression technique that ignores the semantics of data, known as Database Summarization.

In mobile environment, one of the notable constraints is low bandwidth, which is responsible for slow data access. To resolve this issue, [9] proposes Cedar, which effectively manages the weak connectivity by using disk storage and processing power of mobile client. It follows the central organization principal. This approach considers the old copy of a client and the commonality between client and server results. These are helpful to reduce the data volume to be transferred from the database server, hence the communication between client and server becomes more faster while minimizing bandwidth requirement. Moreover, this approach also helps Cedar to solve the problem of replica control in database system across the network.

In a similar vein, [10] proposes an object oriented model to manage mobile database system more effectively. They also addressed the issues related to mobile application development and deployment on a mobile database system by considering GSN based mobile network, which provide data services to allow the accessibility of database server through tiny devices such as portable devices, mobile phones etc. The system also improves the performance due to the access of light weight objects, which are more adoptable by tiny and limited resource-based mobile devices.

One more issue is raised related to the stability of the mobile agents in mobile network due to peer's movement. The free nature of mobile networks in which peers can join and leave the network anytime, provides flexible data hosting as well as improves the data availability into the network by replicating useful data across the network. Hence, as mobile agents are roaming into the network, they may provide better services at the door-step of the mobile users. Additionally, to resolve the issues related to transaction management such as energy consumption, node mobility etc., [12] proposed the strategy to balance the energy consumption on mobile hosts and to reduce the number of failed transactions due to deadline missing.

Moreover, [13] focuses on location-based service oriented mechanisms, in which the characteristics of the information sources have been utilized to do dynamic data management more effectively. They proposed a set of dynamic data management strategies, which reduce the service cost and improve the response time in continuously changing user mobility and data access patterns. To provide easy and fast data access to the user anytime/anywhere in Mobile-Ad-hoc Network (MANET), there is a requirement of processing mobile transactions optimally. There is also an issue of concurrency control in mobile database system. [14] proposes a concurrency control algorithm called Sequential Order with Dynamic Adjustment (SODA), which is based on optimistic concurrency control strategy in mobile environment. To reduce the aborts, the SODA can utilize and adjust the sequential order of committed transactions to improve response time. SODA provides major performance improvement by reducing abort rate.

III. CACHING

To improve the response time of data retrieval in mobile environments, Data caching plays an important role. This section provides an overview of data caching schemes for mobile environments.

A. *Cooperative caching*

In a mobile environment, the mobile client can access data items from the cache of its neighbouring client. This concept is known as “cooperative caching”. Notably, cooperative caching can also be used in conjunction with the P2P paradigm. [15] proposes a cooperative caching scheme, designated as COCA, for mobile systems. COCA categorizes the mobile clients into two categories: Low Activity Mobile clients (LAM) and High Activity Mobile Clients (HAM). Notably, mobile clients from both of these categories share their respective caches. COCA reduces the server workload because the server replicates data items on the LAMs, while the HAMs take advantage of these replicas. Thus, COCA improves the overall system performance, reduces the number of requests as well as the access miss ratio when the mobile hosts are outside of the service region.

Wireless Sensor Networks support several applications such as environment control, intelligent buildings, and target tracking in battlefields. Over the past few years, Wireless Sensor networks have been growing in importance. To serve data in short latency and with minimal energy consumption, these applications require optimization in communication among the sensors. Hence, cooperative data caching protocols has been proposed. The selection of sensor nodes is at the heart of these protocols, and it plays an important role in making the caching and request forwarding decisions. The [16] introduces two new metrics to aid in the selection of such nodes. On the basis of these metrics, the work proposed two new cooperative caching protocols.

B. *Techniques for maintaining cache consistency*

In mobile database systems, if data is cached on a mobile host, it will reduce the query response time and also conserve the generally limited bandwidth. However, there is a need for cache consistency. A basic cache consistency scheme works as follows. The server broadcasts the invalidation report, which identifies the updated data objects so that the mobile hosts may remove the old data from their cache. Due to this reporting, the reconnecting process of a given mobile host may be slow as the mobile host requests the server for validating a cache as it receives an invalidation report. [17] proposes a set of new cache validation schemes, which are capable of conserving the bandwidth for cache validation as well as for query processing.

Caching is also useful for reducing the server load as it facilitates data access at clients, thereby improving the overall performance of the system. In mobile computing environments, there are chances of frequent disconnections. In such situations, coherence between servers and clients becomes a necessity. [18] proposes a category of cache invalidation strategy and mathematical model, and develops a high-performance caching technique. Moreover, the work evaluates the performance for practical wireless mobile computing scenarios.

Furthermore, the cache invalidation methods are record ID based; hence they are not adequate to manage the cache consistency of the mobile clients efficiently. [19] proposes a cache invalidation scheme for continuous partial query in mobile computing environment, which is predicate-based. Here, the cache state of the mobile client is the predicate. The server broadcasts the cache invalidation report (CIR) and the predicate to the client for cache management. This method is useful for reducing the requirement of data for cache management. There are a number of methods to generate the CIR in the server and to identify the invalid data in the client.

Additionally, in dynamic environments, users may not always be able to stay in permanent contact with the network, but message delivery should be guaranteed for all active users of the network. [20] introduces two caching policies: basic caching and leaf caching for providing guaranteed message delivery.

C. *Cache replacement strategies*

While caching frequently accessed data items on the mobile clients improve the system performance, the cache size is generally limited. Hence, effective cache replacement techniques become a necessity to determine the set of data items that should be evicted from the cache. [21] proposes a cache replacement policy called the Weighted Predicted Region-based Cache Replacement Policy (WPRRP) for location-dependent data. WPRRP works on the basis of client’s movement by selecting the predicted region to calculate the weighted distance of a given item.

In a mobile computing environment, the mobile user uses cache to access the data easily, thereby enhancing the data availability as well as improving the data access time. Information is transferred from the server to the query-issuer depending on its current location. This is known as Location

Dependent Information Services (LDISs). [22] proposes a cache replacement policy named Prioritized Predicted Region based Cache Replacement Policy (PPRRP), which uses a cost function for the data eviction based on the client's movement pattern.

[23] proposes a proactive caching model, which caches the result objects along with the index that supports these objects as the results. This is helpful for object reusability for all common types of queries. To optimize the query response time, [23] also proposes an adaptive scheme to cache an index. In mobile environments, proactive caching achieves significant performance gains as compared to page caching and semantic caching.

As the cache size is limited on mobile devices, there are number of cache replacement policies, used to discover a proper subset of items for eviction. The Euclidean distance and Euclidean space are important parameters for eviction in existing policies. In spatial networks, position and movement of the objects are constraints and network distance is an important measure. By considering the network density, network distance and the probability of access, [24] proposes a cache replacement policy which uses Progressive incremental network expansion (PINE) technique to calculate the network distance.

Moreover, [25] proposes a caching policy and broadcast scheme in which the geographical adjacency and characteristics of target area in Location Dependent Queries (LDQ) are reflected. By applying the moving distance of mobile host, [25] develop the caching policy suitable for urban area. The broadcast scheme uses the space-filling curve to cluster data based on adjacency of data in LDQ. The expectation is: when executing LDQ in local cache, the caching policy offers more accurate answers and significantly improves the workload of mobile hosts. Also, the broadcast scheme improves the battery life of the mobile host.

Mobile environment is dynamic, in which the mobile users are moving around a number of service areas. Notably, as the mobile user goes from one service area to another, the new server takes responsibility of that user. This process is known as handoff. In the process of handoff, the new server will not get benefit to access the cache. As a solution to this, [26] discovers numerous cache retrieval schemes to improve the cache retrieval efficiency. The use of 'coordinator buffer' shows the improvement in the cache retrieval. Moreover, Dynamic and Adaptive cache Retrieval scheme (DAR) is developed, which can deal with the service of handoff by utilizing proper cache methods according to specific criteria.

An adaptive per-user per-object cache consistency management (APPCCM) scheme is proposed in [27]. The scheme supports strong data consistency semantics through integrated cache consistency and mobility management in wireless mesh networks. Minimization of overall network cost is the main objective of APPCCM. In APPCCM, caching of data objects is done dynamically, depending on mesh client's mobility and data query/update characteristics and network conditions.

D. Semantic caching

The bandwidth of the mobile devices is also a challenge in developing large spatial database application on mobile environment. Here, the spatial data is used to process the query in mobile environments. [28] attempted to combine multi-resolution spatial data structure and semantic caching techniques for efficient processing of spatial queries. [28] also proposed a new semantic caching model named Multi resolution Semantic Caching (MSC) by considering the characteristics of multi-resolution spatial data and multi-resolution spatial query (MSQ) in mobile environments. MSC improves the performance in three ways: (a) a reduction in the amount and complexity of the remainder queries; (b) the redundant transmission of spatial data already residing in a cache is avoided; (c) a provision for satisfactory answers before 100% query results have been transmitted to the client side.

Furthermore, the two features of semantic cache, namely less network traffic and improved response time, make it efficient for mobile environments. The [29] extends the traditional semantic cache management in three ways: (a) extension of quadtree-based index structures to semantic caches, (b) availability of a query processing strategy and (c) discussion on object-oriented implementation of the semantic cache.

IV. DATA REPLICATION

Replication is the process of sharing information so as to ensure consistency between redundant resources, such as software or hardware components, to improve reliability, fault tolerance, or accessibility. It could be data replication if the same data is stored on multiple storage devices or computation replication if the same computing task is executed many times. A data replication is typically performed in (distributed) space. A computational task is typically replicated in space, i.e. executed on separate devices, or it could be replicated in time, if it is executed repeatedly on a single device. This section discusses the existing data replication approaches to improve data availability in mobile environment.

A. Data Replication in MANETs

A network, where content exchange or delivery is done by autonomous peers, it becomes challenging to construct efficient distributed algorithms for content replication. This is due to the autonomy of the peers and their freedom to decide which objects they want to replicate. Additionally, churn (i.e., peers leaving the network autonomously) poses significant challenges to data availability.

The proposals in [30], [31] discuss replication in MANETs. **E-DCG+** [30] creates groups of mobile peers (MPs) that are bi connected components in a MANET, and shares replicas in larger groups of MPs to provide high stability. An RWR (read write ratio) value in the group of each data item is calculated as a summation of RWR of those data items at each MP in that group. In the order of the RWR values of the group, replicas of items are allocated until memory space of all MPs in the group becomes full. Each replica is allocated at an MP, whose RWR value to the item is the highest among MPs that have free memory space to create it.

The work in [31] aims at classifying different replica consistency levels in a MANET based on application requirements, and proposes protocols to realize them. In this work, each replica is valid till its original owner updates it. Hence, applying strict consistency updates may potentially degrade the system performance, given the inherently dynamic nature of the environment. Thus, the work assumes that all applications do not necessarily require such strict consistency, and it defines consistency based on group-level information consistency. For example, in case of a disaster management group, the information must be consistent within the group, but not strictly consistent w.r.t. to the other groups. Here, the local consistency maintenance within a given group is performed via quorums and it is based on local conditions such as location and time. Notably, the proposals in [30], [31] do not consider an M-P2P architecture and data rarity issues.

Incidentally, P2P replication suitable for mobile environments has been incorporated in systems such as ROAM [32], Clique [33] and Rumor [34]. ROAM, which is a system designed based on the Ward model [35], satisfies a replication solution redesigned specifically for mobile environments. ROAM further considers replication factors such as local replication, appliance compatibility for replication and consistent updates throughout the network.

Clique, a server-less file system model, uses optimistic replication algorithms to store replicas in user's native file systems. It provides mechanisms for ensuring consistent updates (i..e., the replicas are consistent), periodic update management and conflict management. Moreover, it guarantees replica convergence, thereby ensuring data consistency at the group level. In essence, updates are propagated to all nodes within the group to provide reliable and robust data management in the distributed environment.

The Rumor file system is also based on an optimistic replication algorithm, where updates are propagated based on opportunistic cost model among the sites replicating the files. It is built at the application level of the user's mobile devices to provide higher portability, while limiting replication costs. The files are updated through a periodic reconciliation mechanism, which ensures the maintenance of consistency when communication can be restored.

Various data replication techniques have been proposed for MANET databases. By considering the issues of MANET data replication, [36] tries to attempt the classification of existing MANET data replication techniques, and proposes various criteria for selecting the appropriate replication technique for a given application scenario. The work also considers several data replication issues relevant to MANET databases such as energy, mobility, real-time data availability and frequent network partitioning, based on which the replication schemes have been classified.

Moreover, in a MANET, the mobile peers move freely and disconnections take place frequently, thereby reducing the data accessibility due to the dynamically changing network topology. [37] proposes a group mobility model and a replica allocation scheme to address the problem of data accessibility by replicating data items and using the concept of group mobility, where a group of mobile nodes move together.

B. Data Replication in M-P2P Networks

The work in [38] has proposed CLEAR, a context and location-based approach for replica allocation in M-P2P networks. It exploits user mobility patterns, and considers load and different levels of replica consistency.

The works in [39], [40] propose CADRE (Collaborative Allocation and De-allocation of Replicas with Efficiency), which is a dynamic replication scheme for improving the typically low data availability in dedicated and cooperative mobile ad-hoc peer-to-peer (M-P2P) networks. In particular, replica allocation and de-allocation are collaboratively performed in tandem to facilitate effective replication. Such collaboration is facilitated by a hybrid super-peer architecture in which some of the mobile hosts act as the 'gateway nodes' (GNs) in a given region. GNs facilitate both search and replication.

The main contributions of CADRE are as follows. First, it facilitates the prevention of 'thrashing' conditions due to its collaborative replica allocation and de-allocation mechanism. Second, it considers the replication of images at different resolutions to optimize the usage of the generally limited memory space of the mobile hosts (MHs). Third, it addresses fair replica allocation across the MHs. Fourth; it facilitates the optimization of the limited energy resources of MHs during replication.

The proposals in [41], [42] discuss E-ARL, which is a novel Economic scheme for Adaptive Revenue-Load-based dynamic replication of data in dedicated M-P2P networks with the aim of improving data availability. Thus, E-ARL considers a mobile cooperative environment, where the MPs are working towards the same goal, and the network performance is facilitated by the economic scheme. E-ARL essentially allocates replicas based on its economic scheme. Each data item has a price in virtual currency. E-ARL requires a query issuing peer to pay the price of its queried data item to the query-serving peer and a commission to relay peers in the successful query path.

The main contributions of E-ARL follow. First, it uses an economic scheme for efficiently managing M-P2P resources in a context-aware manner by facilitating effective replica hosting and message relaying by peers. Second, it collaboratively performs bid-based replica allocation to facilitate better quality of service. Third, it incorporates both revenue-balancing and load-balancing to improve peer participation and performance. Fourth, it conserves the energy of low-energy MPs to facilitate network connectivity.

The work in [43] considers that M-P2P users may issue queries with varying constraints on query response time, data quality of results and trustworthiness of the data source. Thus, this work proposes ConQuer, which addresses constraint queries in economy based M-P2P networks. ConQuer proposes a broker-based incentive M-P2P model for handling user defined constraint queries. It also provides incentives for MPs to form collaborative peer groups for maximizing data availability and revenues by mutually allocating and de-allocating data items using a royalty-based revenue-sharing method. Such reallocations facilitate MPs in providing better data quality, thereby allowing them to further increase their revenues.

The work in [41] presented the economic model for efficient replica management in M-P2P networks, in which mobile peer has been incentivized to host replica. Here, mobile peers choose which data should be replicated based on its importance. In this manner, mobile peers earn revenues from their hosted queried data items. Hence, it encourages peer participation to improve data availability and discourages free riding. Progressively, [43] proposed ConQuer: a group-based replication method with incentivation in M-P2P networks. This work assume the super-peer architecture for M-P2P network, in which a broker i.e., super-peer has been incentivized for serving constrained query processing by query-issuing peer. Moreover, collaborative peer groups further improves data availability and revenues by mutually allocating and de-allocating data items based on royalty-based model. In a similar vein, a collaborative replication approach for M-P2P networks is also proposed by [40].

The proposal in [44] discussed an economic model LEASE, in which data-providers lease data items to the free-riders in lieu of a lease payment. Hence, it provides free-riders the opportunity to earn revenue by hosting data, thereby incentivizing them towards data hosting. [45] also discussed incentive-based services for a dynamic data management in M-P2P networks.

V. MESSAGE COMMUNICATION

A. *Ad hoc Network*

A network without any base station is known as Ad hoc Network also known as infrastructure less or multi-hop network, where every node is a mobile which facilitates peer-to-peer and peer-to-remote communication any time anywhere.

A new generation of mobile application is promoted by the advancement in mobile phone technologies which facilitate users to interact with one another and information sharing when the mobile devices are physically near. Such applications take advantage of ad-hoc network to provide location based services and use the ad-hoc network a as filtering mechanism.

The ad hoc network with replicas is constrained by the consistency management of data operations. The work in [46] proposes two management consistency protocols by considering the special types of applications in which each data item can be partitioned. The results of simulations give idea about the characteristics of these protocols in mobile ad hoc network. These protocols maintain the global consistency in the entire network by improving the data availability and reduce the traffic for data operations.

The difficult task in multi-hop communication is frequently changed network topology by mobile nodes in wireless ad hoc networks. By targeting the various scenarios of design space, the wireless community designed hundreds of new routing protocols in the last 15 years. The objective of [47] is to create taxonomy of the ad hoc routing protocols, and to survey and compare representative examples for each class of protocols.

The Mobile Ad Hoc NETworks (MANETs) may facilitate the communication between the group members in distributed wireless network environment (Group Communication). There are number of group-oriented applications available like disaster management, battlefields, e-commerce, e-education, etc. The dynamic construction of efficient and reliable multicast routes and variety of channels in context of user mobility are the demand of the group communications. By considering various performance measures such as energy efficient route establishment, packet delivery ratio, quicker and faster proactive route recovery, network life time, reliability, Quality of Service (QoS) based on bandwidth, delays, jitters, and security; the researchers improved the Multicast routing mechanism in MANETs.

The review presented in [48] focuses on multicast routing mechanisms with respect to different topological routing categories that facilitate the multimedia communication over MANETs. Moreover, provides future directions for the research and development.

B. *Geological*

In mobile computing environment, each node is a mobile device, i.e., moving object means there is no concept of fixed infrastructure and no central access to the data; this is the key feature behind the development of the several data access methods like, access data related to a particular location, access data from the nearest neighbour, etc.

The replication and forwarding techniques caused by mobility make data availability within the region of interest. This work in [49] proposes a decentralized algorithm for data dissemination by associating data with geo-address, which describe the Point of Interest (POI) of data. In this case the purpose of replication of data is to increase data availability within circular area around the POI.

The Sector Heads Aided Flooding Technique (SHAFT) is used to avoid unnecessary communication overhead that caused by the mobile nodes that forward data by arranging data placement geometrically in the area. The locally measured density of mobile devices in range is to be adopted by the algorithm. The usefulness of the approach is shown practically by applying to a cooperative parking lot management system based on Manhattan mobility model.

The distributed wireless mobile network system provides infrastructure and important applications to the digital ecosystems one of them is mobile navigations and continuous mobile information services. Generally in mobile environment, the query services are working by finding the Point of Interest or interest objects while mobile users are moving; these services are continuous query processing or continuous k nearest neighbour (CKNN). The solution of continuous k nearest neighbour (CKNN) search by proposing a neural network based algorithm which divides the query path into segments and improves the overall query process is mentioned in [49].

In mobile computing environment generally all the operations are based on location because in this environment all the nodes are mobile and not bound within fixed infrastructure; in this case the location based queries are gradually more in demand and one of the notable reasons behind this is the wide availability of mobile devices. Range searching method is also used frequently which work inside pre-defined area by returning the objects of interest. Number of existing methods is based on road network expansion method in which the nodes are expanded using intersections and distance from node to the query point is calculated but this approach is inefficient and extremely complex. On the basis of

Voronoi diagram, two new methods namely Voronoi Range Search (VRS) and Voronoi Continuous Range (VCR) are proposed in [50]. In VRS, road networks are partitioned into some polygons for accurate and efficient process of range search queries. VCR proposed based on VRS which is particularly used for moving queries (Continuous Range Search Queries).

C. Vehicular Network (VANET)

VANET is one type of mobile ad hoc network where each vehicle act as a router either V2V i.e., Vehicle to Vehicle communication or V2I i.e., Vehicle to Infrastructure communication take place. This simply means, *Network on wheels*.

To perform agent migration and mobility among nodes (handover) in an infrastructure less vehicular ad hoc network (VANET) the concept of opportunistic communication is evaluated and by considering this, [51] looks at location awareness to support agent mobility. One of the interesting applications of this idea is "virtual sensor network" composed of software agents that done the task in form of sensing services with the help of available resources provided by physical nodes like computer platforms, communication devices or physical sensor devices.

Exchanging events in vehicular ad hoc network (VANET) helps drivers to receive information about relevant places or to avoid undesirable situations. The work in [52] presents data management as a solution for event exchange in vehicular networks and comparison of two different approaches. From that the first approach is helpful to estimate the relevance of events by calculating of geographic vectors; and second approach exploits digital road maps. For the testing of proposals in a real environment, a prototype is described here. The usefulness of utilizing the information stored in digital road maps for data management and sharing in vehicular networks as well as the first approach is useful in cars where maps are not available has been proved by an exhaustive simulation-based experimental evaluation.

Moreover, the spatial queries and mobile database queries become hot spot issues because the explosion of Geographic Information System (GIS) and Global Positioning System (GPS), classically, range search and k nearest neighbour queries. The study from Euclidean distance to network distance and from dealing with static objects to moving objects shows the development in this area from many years. Only few literature reports have dealt with Continuous Range Search (CRS) query processing. In CRS the query path needs to be segmented that's why some existing approach to CRS is still impractical [53].

VI. CROWDSOURCING

Crowdsourcing refers to the process of outsourcing activities from a firm to an online community or crowd in the form of an 'open call'. Any member of the crowd can then complete an assigned task and be paid for their efforts [54]. In this concept the company pays only for products or services that meet its prospects; beyond the cost and benefits of the company. Characterization of crowdsourcing from management science perspective is given in [55].

The task that requires human intelligence for the crowdsourcing has become an acceptable medium for creation of resources. For the purpose of system building and evaluation, the information retrieval and related fields regularly use it. In this case there are chances of fraudulent attempts by malicious workers and it is also challenging and time consuming process to identify these persons for both crowdsourcing providers and requesters. The work in [56] explains that how to reduce such fraud attempts.

The usage of mobile devices is going to increase day by day because of the availability of number of facilities other than just a communication device. Mcleark, AMT, txteagle, mCrowd and SMSAssassin are the various applications and platforms developed to take the best advantage of the crowdsourcing. Moreover, few mobile companies are planning to provide crowdsourcing enabled mobile device to support such applications.

Moreover, [57] presents a system to provide open-access working platform for people who can earn small amount by completing the tasks e.g., translations, transcriptions, surveys etc. Here, people are paid either in airtime or MPESA (mobile money) by the corporations. Such services have been recently launched in few countries like Kenya. When Crowdsourcing is extended to sensor-rich mobile devices like smart phones, it has potential that can be truly set free. The work in [58] proposes a new iPhone-based mobile crowdsourcing platform called mCrowd; which facilitates users to work on sensor related crowdsourcing tasks at fingertips e.g., geolocation aware image collection, image tagging, road traffic monitoring etc. through the rich sensor equipped with iPhone.

We can relate or combine the concept of crowdsourcing with mobile for the effective usage of mobile devices and for the beneficial usage of the Crowdsourcing; it is not just enough we can also relate the crowdsourcing with database systems. Sometimes the queries processing requires human input as they cannot be answered by machines only. For ex: the queries like matching, ranking or aggregating results based on fuzzy criteria. To process the queries, which is sufficiently answered neither by database system nor does search engine require human input via crowdsourcing which is used by CrowdDB. The CrowdDB uses the SQL for two purposes: as a language for posing complex queries and as a way to model data. There are some differences between CrowdDB and traditional database system like CrowdDB waits for the human inputs and performance and cost of the query depend on a number of new factors. The [59] describes the design of CrowdDB.

VII. CONCLUSION

This survey presents various mobile applications, including facility of data sharing and data storing for easy and handy access to the static and dynamic information in the mobile databases. It also discussed the various communication strategies in mobile ad hoc networks. Furthermore, the survey also explores various caching mechanisms and replication schemes to improve respectively response time and data availability in mobile environment. Finally, the survey also discussed about the crowdsourcing, which is an upcoming mobile data management field. In future, we will discuss the important research issues in more detail related to mobile databases.

REFERENCES

[1] International Telecommunication Union, http://www.itu.int/ITUD/ ict/statistics/.

[2] N. Tao, K. Zhu, and X. Fu, "Thumb: A real-time resource information sharing application over mobile phones," in In Proc. of IEEE Conference on (PERCOM Workshops), 2012, pp. 528–530.

[3] S. Nasution, E. Husni, and A. Wuryandari, "Prototype of train ticketing application using Near Field communication (NFC) technology on Android device," in In Proc. of Sys. Engi. and Tech. (ICSET), 2012, pp. 1–6.

[4] M. Ramesh, T. Anu, and H. Thirugnanam, "An intelligent decision support system for enhancing an m-health application," in In Proc. of (WOCN), 2012, pp. 1–5.

[5] M. Ramesh, S. Anand, and P. Rekha, "Mobile Software Platform for Rural Health Enhancement," in Advances in Mobile Network, Communication and its Applications (MNCAPPS), 2012 International Conference on In Proc. of MNCAPPS, 2012, 2012, pp. 131–134.

[6] V. Kumar, Mobile Database System. Wiley - Interscience.

[7] D. Barbará, "Mobile Computing and Databases-A Survey," IEEE TKDE, vol. 11, no. 1, pp. 108–117, 1999.

[8] D. Chan and J. F. Roddick, "Context-sensitive mobile database summarisation," Proc. ACSC, pp. 139–149, 2003.

[9] N. Tolia, M. Satyanarayanan, and A. Wolbach, "Improving mobile database access over wide-area networks without degrading consistency." ACM, 2007, pp. 71–84.

[10] K. A. Qureshi, S. Mohiuddin, A.-U. Aziz-Uddin, and A.-U.-R. Atique- Ur-Rehman, "A strip-down database for modern information systems," Proc. Computers, pp. 81–88, 2010.

[11] S. K. Madria, M. Baseer, V. Kumar, and S. Bhowmick, "A transaction model and multiversion concurrency control for mobile database systems," Distrib. Parallel Databases, vol. 22, no. 2-3, pp. 165–196, 2007.

[12] L. Gruenwald, S. M. Banik, and C. N. Lau, "Managing real-time database transactions in mobile ad-hoc networks," Distrib. Parallel Databases, vol. 22, no. 1, pp. 27–54, 2007.

[13] S.-Y. Wu and K.-T. Wu, "Effective location based services with dynamic data management in mobile environments," Wireless Networks, vol. 12, no. 3, pp. 369–381, 2006.

[14] Z. Xing, L. Gruenwald, and S. Song, "An optimistic concurrency control algorithm for mobile ad-hoc network databases." ACM, 2010, pp. 199– 204.

[15] C.-Y. Chow, H. V. Leong, and A. Chan, "Peer-to-Peer Cooperative Caching in Mobile Environments." IEEE Computer Society, 2004, pp. 528–533.

[16] N. Dimokas, D. Katsaros, L. Tassiulas, and Y. Manolopoulos, "High performance, low complexity cooperative caching for wireless sensor networks," Wireless Networks, vol. 17, no. 3, pp. 717–737, 2011.

[17] H. Kang and S. Lim, "Bandwidth-Conserving Cache Validation Schemes in a Moblie Database System." Springer-Verlag, 2001, pp. 121–132.

[18] J. Zhang, Y. Chu, and J. Yang, "A category on the cache invalidation for wireless mobile environments." Springer-Verlag, 2006, pp. 939–942.

[19] Y. D. Chung, "A cache invalidation scheme for continuous partial match queries in mobile computing environments," Distributed and Parallel Databases, vol. 23, no. 3, pp. 207–234, 2008.

[20] V. Sourlas, G. S. Paschos, P. Flegkas, and L. Tassiulas, "Caching in content-based publish/subscribe systems." IEEE Press, 2009, pp. 1401–1406.

[21] A. Kumar, M. Misra, and A. K. Sarje, "A weighted cache replacement policy for location dependent data in mobile environments," Proc. Symposium on Applied Computing, pp. 920–924, 2007.

[22] A. Kumar, A. K. Sarje, and M. Misra, "Prioritised Predicted Region based Cache Replacement Policy for location dependent data in mobile environment," International Journal of Ad Hoc and Ubiquitous Computing, vol. 5, no. 1, pp. 56–67, 2010.

[23] H. Hu, J. Xu, W. S. Wong, B. Zheng, D. L. Lee, and W.-C. Lee, "Proactive Caching for Spatial Queries in Mobile Environments," Proc. of Data Engineering, pp. 403–414, 2005.

[24] M. M. Janef, R. Parameswaran, R. Nadarajan, and M. Safar, "PINEguided cache replacement policy for location-dependent data in mobile environment," Proc. of Pervasive Tech. Related to Assistive Environments, pp. 16:1–16:5, 2008.

[25] I.-d. Jung, Y.-h. You, J.-h. Lee, and K. Kim, "Broadcasting and caching policies for location-dependent queries in urban areas." ACM, 2002, pp. 54–60.

[26] W.-C. Peng and M.-S. Chen, "Design and Performance Studies of an Adaptive Cache Retrieval Scheme in a Mobile Computing Environment," IEEE TMC, vol. 4, no. 1, pp. 29–40, 2005.

[27] Y. Li and I.-R. Chen, "Adaptive per-user per-object cache consistency management for mobile data access in wireless mesh networks," Journal of Parallel and Distributed Computing, vol. 71, no. 7, pp. 1034–1046, 2011.

[28] S. Sun, X. Zhou, and H. T. Shen, "Semantic caching for multiresolution spatial query processing in mobile environments," Proc. of SSTD, pp. 382–399, 2005.

[29] D. Li, M. Hu, and R. Chen, "Quadtree-based management on semantic cache for mobile computing," International Journal of Computer Applications in Technology, vol. 43, no. 4, pp. 393–399, 2012.

[30] T. Hara and S. K. Madria, "Data replication for improving data accessibility in ad hoc networks." IEEE TMC, vol. 5, no. 11, 2006.

[31] T. Hara and S. Madria, "Consistency management among replicas in peer-to-peer mobile ad hoc networks." Proc. IEEE SRDS, 2005.

[32] D. Ratner, P. Reiher, G. Popek, and G. Kuenning, "Replication Requirements in Mobile Environments," Mobile Networks and Applications, vol. 6, no. 6, 2001.

[33] B. Richard, D. Nioclais, and D. Chalon, "Clique: A transparent, peerto- peer replicated file system." Proc. MDM, 2003.

[34] R. Guy, P. Reiher, D. Ratner, M. Gunter, W. Ma, and G. Popek, "Rumor: Mobile data access through optimistic peer-to-peer replication." Proc. ER Workshops, 1998.

[35] D. Ratner, G. J. Popek, and P. Reiher, "The Ward Model: A Scalable Replication Architecture for Mobility," In Workshop on Object Replication and Mobile Computing, 1996.

[36] P. Padmanabhan, L. Gruenwald, A. Vallur, and M. Atiquzzaman, "A survey of data replication techniques for mobile ad hoc network databases," The VLDB Journal, vol. 17, no. 5, pp. 1143–1164, 2008.

[37] J.-L. Huang and M.-S. Chen, "On the Effect of Group Mobility to Data Replication in Ad Hoc Networks," IEEE TMC, vol. 5, no. 5, pp. 492– 507, 2006.

[38] A. Mondal, S. K. Madria, and M. Kitsuregawa, "CLEAR: An Efficient Context and Location-Based Dynamic Replication Scheme for Mobile- P2P Networks," Proc. DEXA, pp. 399–408, 2006.

[39] A. Mondal, S. K. Madria, and M. Kitsuregawa, "CADRE: A Collaborative replica allocation and deallocation approach for Mobile-P2P networks," Proc. IDEAS, pp. 21–28, 2006.

[40] A. Mondal, S. K. Madria, and M. Kitsuregawa, "A Collaborative Replication Approach for Mobile-P2P Networks," IJHCR, vol. 1, no. 2, pp. 75–97, 2010.

[41] A. Mondal, S. Madria, and M. Kitsuregawa, "EcoRep: An Economic Model for efficient dynamic replication in Mobile-P2P networks," Proc. COMAD, pp. 116–127, 2006.

[42] A. Mondal, S. K. Madria, and M. Kitsuregawa, "E-ARL: An Economic incentive scheme for Adaptive Revenue-Load-based dynamic replication of data in Mobile-P2P networks," Distributed and Parallel Databases, vol. 28, no. 1, pp. 1–31, 2010.

[43] A. Mondal, S. Madria, and M. Kitsuregawa, "ConQuer: A Peer Group- Based Incentive Model for Constraint Querying in Mobile-P2P Networks," Proc. MDM, pp. 134 –139, May 2007.

[44] A. Mondal, S. K. Madria, and M. Kitsuregawa, "LEASE: An Economic Approach to Leasing Data Items in Mobile-P2P Networks to Improve Data Availability," Proc. DNIS, pp. 222–231, 2007.

[45] S. K. Madria and A. Mondal, "Economic-based Incentive Schemes for Dynamic Data Management in Mobile P2P Computing," Proc. MDM, p. 228, 2008.

[46] T. Hara and S. Nishio, "Global consistency management methods based on escrow approaches in mobile ad hoc networks," Mobile Information Systems, vol. 6, no. 1, pp. 5–26, 2010.

[47] A. Boukerche, B. Turgut, N. Aydin, M. Z. Ahmad, L. Bölöni, and D. Turgut, "Survey Paper: Routing protocols in ad hoc networks: A survey," Computer Networks, vol. 55, no. 13, pp. 3032–3080, 2011.

[48] R. C. Biradar and S. S. Manvi, "Review: Review of multicast routing mechanisms in mobile ad hoc networks," Journal of Network and Computer Applications, vol. 35, no. 1, pp. 221–239, 2012.

[49] H. Meyer and K. A. Hummel, "A geo-location based opportunistic data dissemination approach for MANETs," Proc. CHANTS, pp. 1–8, 2009.

[50] K. Ahmadian, M. Gavrilova, and D. Taniar, "Multi-criteria optimization in GIS: continuous k-nearest neighbor search in mobile navigation," Proc. of Computational Science and Its Apps, pp. 574–589, 2010.

[51] E. P. De Freitas, T. Heimfarth, F. R. Wagner, A. M. Ferreira, C. E. Pereira, and T. Larsson, "Geo-aware handover of mission agents using opportunistic communication in VANET," Proc. ruSMART/NEW2AN, pp. 365–376, 2010.

[52] T. Delot, S. Ilarri, N. Cenerario, and T. Hien, "Event sharing in vehicular networks using geographic vectors and maps," Mobile Information Systems, vol. 7, no. 1, pp. 21–44, 2011.

[53] K. Xuan, G. Zhao, D. Taniar, B. Srinivasan, M. Safar, and M. Gavrilova, "Continuous range search based on network Voronoi diagram," International Journal of Grid and Utility Computing, vol. 1, no. 4, pp. 328–335, 2009.

[54] P. Whitla, "Crowdsourcing and Its Application in Marketing Activities," Contemporary Management Research, vol. 5, pp. 15–28, 2009.

[55] C. G. Eric SCHENK, "Crowdsourcing: What can be Outsourced to the Crowd, and Why ?" 2009.

[56] A. P. d. V. Carsten Eickhoff, "How Crowdsourcable is Your Task?" Proc. CSDM, February,9 2011.

[57] N. Eagle1, "txteagle: Mobile Crowdsourcing," Lecture Notes in Computer Science, vol. 5623, pp. 447–456, 2009.

[58] R. H. D. G. M. C. Tingxin Yan, Matt Marzilli, "Demo Abstract: mCrowd - A Platform for Mobile Crowdsourcing," ACM 978-1-60558-748-6, 2009.

[59] M. J. Franklin, D. Kossmann, T. Kraska, S. Ramesh, and R. Xin, "CrowdDB: answering queries with crowdsourcing," ser. SIGMOD '11. New York, NY, USA: ACM, 2011, pp. 61–72.