Sarddar *et al. Braz J Sci Technol (2015) 2:3
DOI 10.1186/s40552-015-0008-x

Brazilian Journal of Science and Technology
a SpringerOpen Journal

RESEARCH Open Access

A novel approach on weight based optimized routing for mobile cloud computing

Debabrata Sarddar[†], Rajesh Bose[†] and Sudipta Sahana[*†]

*Correspondence: ss.jisce@gmail.com
[†]Debabrata Sarddar, Rajesh Bose and Sudipta Sahana contributed equally to this work
Department of CSE, University of Kalyani, Nadia, West Bengal, India

Abstract
The combination of mobile communications and cloud computing are growing immensely popular. Application designers and architects use cloud technology advantages to deliver data and information to mobile devices as swiftly and seamlessly as possible. However, a critical aspect in this scenario is the factor of routing. Complex routing techniques can use up a lot of energy. Coupled with an increasing number of cloud data centers, this has the potential to leave a negative impact on our environment. An efficient routing model can help lower carbon emissions in the long run. Most existing routing methods involve shortest path routing to ease congestion in mobile networks. Through this paper, the authors examine the shortcomings of such existing protocols. The authors discuss mobile base stations and how routing takes place. An algorithm is presented which illustrate the application of the proposed model as presented in this paper. The proposed model is introduced by the authors in this work as Weight-Based Optimized Routing. Using this new approach, the authors demonstrate a method to lower network congestion and ensure a more smooth data exchange between mobile devices and cloud nodes. A reduction in mobile network congestion situations and enhancement of data transmission rates are vital towards achieving a green cloud computing environment.

Keywords: Cloud computing, Mobile computing, Routing, Virtualization, Base station

Background

Mobile cloud computing (MCC) is a combination of cloud computing, mobile telephony and wireless networks. It is essentially a concept which allows streaming of feature rich mobile applications in a manner that is beneficial for mobile device users, service and application providers. Mobile device users no longer have to bank on expensive hardware in order to experience quality applications that might require either extensive computing performance beyond the existing capacities offered by current hardware. Service providers can, at the same time, scale their operational requirements based on subscriber base. In other words, service providers can house their operations in Infrastructure-as-a-Service cloud computing models. Application developers, similarly, can hire Platform-as-a-Service services to develop, test and market their products. In a Software-as-a-Service cloud model, entire software suites and products can be served to users over internet and mobile networks without the need for installing whole software on users' computing devices. The pyramid-like structure of how these cloud service models are layered one after the other is depicted in Figure 1.

Figure 1 Types of cloud.

In a MCC environment, mobile users are able to enjoy software and high quality video streaming without the need to burden the hardware resources of their mobile devices. In an environment such as this, mobile networks and cloud computing system interact with each other in order to offer users an experience that is progressively richer at optimized costs. For example, a mobile game with an engaging user interface might use up memory and CPU resources of a mobile very quickly, thus, leaving very little for other applications to operate on. Situations such as this can cause a compromise that most users or resident applications within a mobile device, may reject outright.

With the aid of cloud computing, the bulk of data processing activities and storage of data can be moved away, thereby, minimizing hardware upgrade requirements which could otherwise be costly. MCC environments also allow an opportunity of reducing heat emissions from mobile devices. Limiting computational loads on mobile device processors leads to lower power consumption, too. In this paper, we discuss some of existing research works on mobile routing. We explain how an effective means of routing can ease inter-network congestion. We also demonstrate a routing algorithm that can be used to implement effective communication in cases where a mobile device may be at the intersection of two mobile base stations. Routing closely resembles bridging. The authors through this research paper, show how an effective routing can not only imply the shortest distance between two points of communication, but also the shortest path by which energy consumption at the base stations and intermediary nodes can be optimized.

The rest of this paper has been sub-divided into seven primary sections. We discuss related works and present our literature survey in "Related works". In "Method", we present the workings of our proposed method. Our proposed algorithm has been explained in "Algorithm". The flowchart to our proposed method has been illustrated in "Flow chart for the proposed method". In "Results and discussion", we discuss the results and conduct an examination on our observations. Finally, we conclude our work in "Conclusion".

Related works

In their paper "A Mobile Cloud Computing Architecture with Easy Resource Sharing", Sarddar et al. (2014) have proposed a model which would allow mobile device users to use cloud-based applications. Through their research paper, they have demonstrated use of a resource manager that enables easy resource sharing. Not only does this translates into savings in terms of power consumption, but also shortens the time it takes for communications to take place.

Discussing frameworks for execution of mobile applications in their paper titled "Application optimization in mobile cloud computing: Motivation, taxonomies, and open challenges", the authors, Ahmed et al. (2015) have shed light on optimization strategies that can provide enhance performance of mobile applications in mobile cloud computing environments. The authors have also expounded on several mobile application execution frameworks and their effects on mobile applications and devices in a mobile cloud computing environment. The work features an in-depth comparison of application frameworks, and investigates suitable metrics to boost execution performance of mobile applications. The authors have also suggested the need for research to be made in designing frameworks considering that their findings point to certain flaws in the existing frameworks. According to them, some of the frameworks improve response times by reducing delays over WAN. In their opinion, a well-designed application framework should be able to offer efficiency without involving extensive programmer support or intensive deployment of virtual machines.

Determining the availability of a mobile device is measured by mobility factor. In their work presented by Park et al. (2012), a high mobility aspect was found to have a reliability issues owing to loss of connection. In the paper presented by the authors of the work, users have been categorized into groups considering the mobility factors involved. The groups have been classified into low mobility, middle mobility, and high mobility. The classifications are used to calculate the probability of mobility. However, the technique may be considered infeasible in cases where determination of mobility becomes a challenge in open public spaces. This approach is workable in situations where the environment is a closed bounded area such as office buildings, colleges, libraries, etc.

A grid model is discussed by Deboosere et al. (2011) in their work in which the authors discuss mobile device as a thin client. In their paper, the authors of this work have shown how a mobile connects to servers using thin clients like VNC (Virtual Network Computing) or streaming protocols. Using such a system, user inputs are sent through wireless network to a designated server. After processing of the input, the server returns a graphical output to the mobile device. Their research focuses on server selection algorithms which are relevant in situations where a mobile device changes its geographical locations. To keep transmissions delays to a minimum, it may be necessary for the application involved to be shifted to a nearby server. As this might result in affecting the performance, the authors have aimed at providing effective algorithms to mitigate the effects while leaving support for operator mobility intact.

Chun et al. (2011) have preferred migration of virtual machines through 3G or Wi-Fi networks. Since, application assignment to a resource server is conducted using VM migration, creation of subsequent resultant device clones ensure that mobile applications remain unchanged and, further, there is no need to explain the procedures involved in the work presented by Cuervo et al. (2010). The authors have also presented a "cost model" that shows the cost involved in cloud migration and implementation. In doing so, they have associated such cost vis-à-vis a uniform implementation.

While directing their attentions on architecting a planning model for transmission of data to the cloud that is cost-effective, Cho et al. (2010) have designed Pandora. Later, these authors have gone on to propose a solution to curb transmission latency under budget constraints (Cho and Gupta 2011). Their study differs from ours in that they lay emphasis on fixed volume data transfer which is static, rather than data that is produced dynamically. Further, they have considered a single-cloud measure, whereas our study involves more than one data hub.

Virtual machine migration and code partitioning comes together in the amalgam used by Maui et al. (2010). The primary aim of the authors is to save energy. Applications are separated from mobile phones and moved to local and remote servers. This split is performed during execution and is designed to by dynamic. For connectivity, their work has accommodated use of either 3G or Wi-Fi networks. Developers document an application's methods which can be offloaded at implementation time. In case a remote server can be accessed, the model determines whether to proceed with off-loading.

In a paper written by Marshall et al. (2010), the authors have presented a cloud model that is elastic and is able to respond to the demands of an application. At the core of this model, an elastic site manager controls resource provisioning. The paper contains detailed evaluations of the proposed model, and offers explanations with regard to physical cluster and cloud resources.

While estimating the performance of cloud through application of Gang scheduling algorithms, Moschakis and Karatza (2012) have also discussed frequent communication in the context of which such algorithms have been deemed appropriate. Their study involves only a single public cloud consisting of a cluster of virtual machines on which parallel jobs are processed.

A cloudlet concept is an idea that has been expounded by Satyanarayanan et al. (2009) in their research in mobile cloud computing. The authors harness virtual machine technology to deliver services across mobile and Wi-Fi networks as quickly and effectively as possible. In their work, cloudlets appear as either a computer that is sufficiently equipped and quite capable to handle resource and computing demands or as a cluster of computers similarly equipped, and which is/are connected to the internet and ready for use by mobile devices in proximity.

Mobile cloud computing is influenced by concepts which are similar to a simple cloud computing framework. Mei et al. (2008) in their paper discusses the need for certain requirements that are to be met which are common in case of both MCC as well as in general cloud computing environments. These factors or requirements are adaptability, scalability, availability and self-awareness.

Of late, considerable research has been made towards reducing consumption of energy in telecommunications networks. The works of authors (Shen and Tucker 2009; Cavdar

2011) have focused on optical core networks in this regard. However, only select works have, thus far, made any significant contribution in attempting to build solutions that are able to slash CO_2 emissions in the operation and overall framework of an optical network. This is markedly a unique target vis-à-vis that of reducing overall energy consumption; more so, when renewable energies are taken into consideration.

In order to reduce CO_2 in networks those are of the type "IP over WDM" (either with data centers or without), the authors in (Dong et al. 2011) propose a method by making use of renewable networks. They have demonstrated a Linear Programming (LP) model for a network design that cuts back on CO_2 emissions (a low carbon design) and propose a heuristic approach to enhance utilization of renewable energy.

Where optical devices can draw power either from legacy energy sources or renewable ones, the contribution of the authors (Ricciardi et al. 2011) gains traction. In their work, the authors have proposed an approach that works through formulation and comparison of energy-aware static RWA strategies for WDM networks.

Across the world, various research projects over the years have conducted investigations on how to effectively engage renewable energies in the telecom network infrastructure.

Method

Our proposed method is primarily directed at proper route selection at mobile network level for message delivery. As is well recognized, routing forms the crux of mobile communication. Improper routing can set off problems such as delays, jitters, packet drops, etc. in a network. In this paper, we have proposed a weight-based route selection algorithm which works off of two fundamental network properties involving cost of path and channel capacity of a link. Our proposed algorithm, also takes into account selection of base station before routing calculation and implementation can take place.

In a situation where a mobile device within an area marked as an intersection of two base stations, it is becomes necessary to identify the particular base station to which the mobile device needs to connect. To arrive at a solution, we have identified two network factors comprising utility factor and signal strength for a given base station in that particular area. A base station that has a larger weight factor value (W1) will be selected as the home base station. The default base station will be considered to be as the home base station when a mobile device is not in any intersection.

In our proposed model, each base station would hold a small device named link manager (LM). The LM would communicate with Mobile Switching Center to assess and seek out various links for transmission between source and resource manager (RM) (Sarddar et al. 2014). LM calculates the weight factor (W2) for different link subject to path cost and channel capacity. The link with the greater weightage factor is selected as the optimized route.

As might be obvious, in cases where a link is sufficiently free to handle traffic, the shortest path is selected according to W2. The RM acts as a central database repository for all cloud nodes registered with it. It is responsible for periodically updating the routing table. Transmitting message from the mobile device to the appropriate resource is the responsibility of the RM. Using the proposed method, incidences of network congestion can be significantly brought down. Each and every link can also be sufficiently utilized. Optimization of time also takes place as a result.

Calculation of W1

The weight factor W1 is a function of utility and signal strength of a base station. At a given particular location, a mobile device is able to compute signal strength of stations that it can reach. Utility is expressed in terms of response provided in a single unit of time (throughput) and the response capacity of a given station.

Throughput calculation is given as follows:

$$\text{Throughput} = \frac{\text{Response}}{\text{Unit of time}}$$

Thus, utility of a base station is given as follows:

$$\text{Utility} = \frac{\text{Throughput}}{\text{Maximum capacity}}$$

Therefore, Weight Factor (W1) can be computed as:

$$\text{Weight factor (W1)} = \frac{\text{a} \times \text{utility} + \text{b} \times \text{signal strength}}{\text{a} + \text{b}}$$

("a" and "b" are separate weight factors; where a > b).

Base station with the higher W1 value will be selected as the home base station.

Calculation of W2

Weight factor W2 is dependent on path cost and channel capacity. Path cost corresponds to hop count when negotiating the route from source to destination.

Channel capacity for a different route (C) is given as:

$$\sum \frac{(c_1, c_2, \ldots, c_n}{n}$$

Where c1, c2, c3 indicate channel capacities for each link between two consecutive hops.

Weight factor "W2" is computed as:

$$\text{Weight factor (W2)} = \frac{\text{c} \times \text{path cost} + \text{d} \times \text{channel capacity}}{\text{c} + \text{d}}$$

("c" and "d" are different weight factors where d > c).

Route with the maximum W2 value is chosen for data communication.

Algorithm

The algorithm of our proposed model is given as follows:

1. A mobile device begins checking its own location.
2. If it finds its location to be in an intersection formed by two base stations, W1 is called.
3. If otherwise, the only base station detected is set as the home base station.
4. The mobile device transmits a connect request to the base station.

5. The resident LM in that base station sends to the MSC a route request message.
6. The MSC responds with details of all possible routes leading up to the RM.
7. An optimized route selection process using W2 method is invoked by LM.
8. After computation, the base station feeds the mobile device the optimized route calculated.
9. The mobile device transmits the data packet to the RM through the optimized route.
10. In turn, the RM checks the node address of the destination location and begins consulting its own routing table.
11. Upon locating the appropriate entry, the RM forwards the packet to the final destination and returns an acknowledgement to the mobile device.
12. The link is terminated when communication is over.

Variable description

BS Base station
MD Mobile device
MSC Mobile switching Center
RM Resource manager
LM Link manager
WI Weight factor regarding base station
W2 Weight factor regarding routing

The mobile cloud computing and networking structure is presented in Figure 2. The flowchart for our proposed method is produced below.

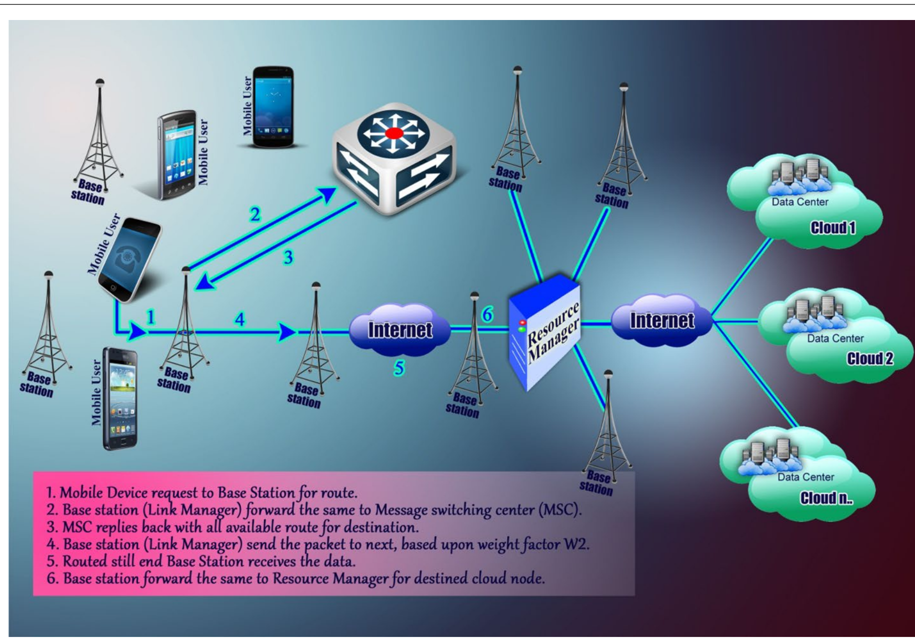
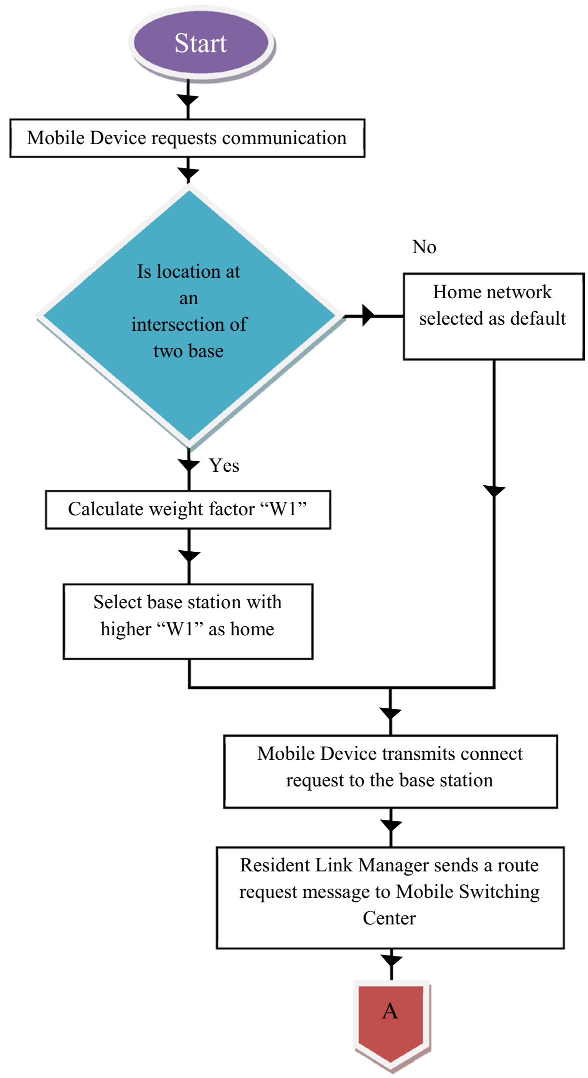
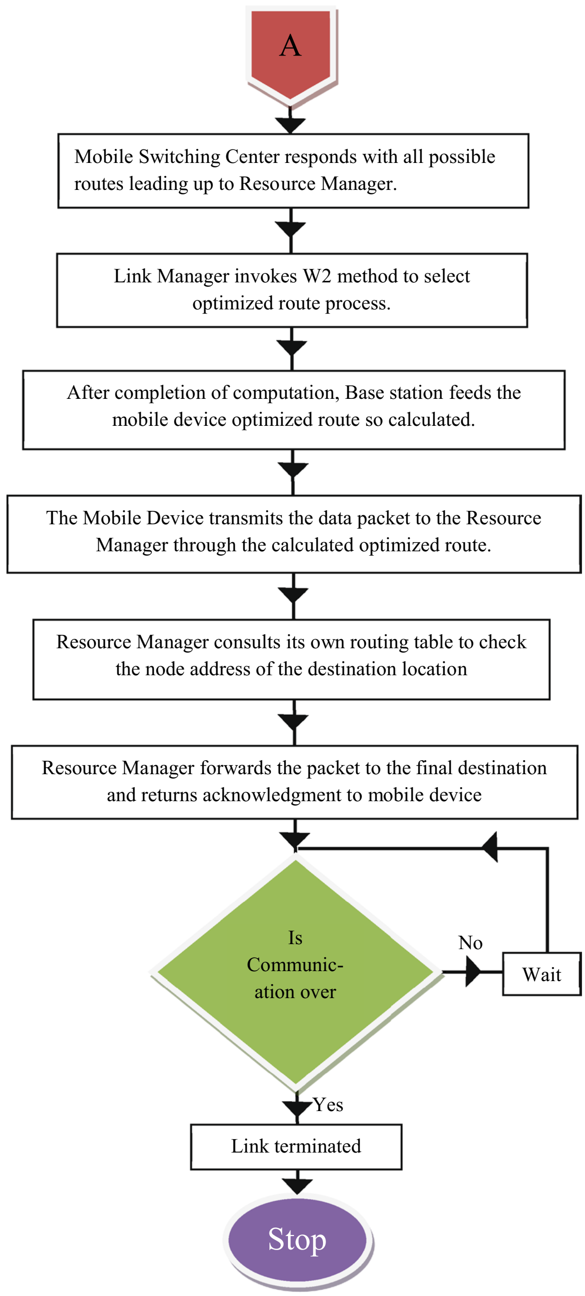


Figure 2 Proposed network architecture.

Flow chart for the proposed method





Results and discussion

It has been observed that many authors have directed their efforts and attention on capitalizing on shortest path routes for rationalizing time and energy utilization. A brief literature survey conducted reveals that this approach finds favor among many researchers even though a route may be found to be congested.

Suppose that there are three different ways to communicate between router A and B. Assuming that each link is capable of handling 50 concurrent communications, any additional request for connection over and above 50 would cause an overload that would begin emitting more carbon emissions. New connection requests arriving at this node would have to wait fuelling the demand for energy as well as time. The shortest path algorithm is shackled by this problem. We use weight factors for routing path selection in our proposed Weight-Based Optimized Routing for Mobile Cloud Computing. This approach not only reduces communication link overheads as can be seen upon observing and comparing with the more traditional shortest path model, but also reaches a point where energy and time optimization can be said to have been achieved. Our proposed model cuts down delays in queues. This, in turn, reduces CO_2 emissions, and is a step forward in the direction of green cloud computing. With the aid of our proposed model, all available communication links can be utilized in an optimized manner.

Simultaneous with the result analysis depicted in the graph, we have demonstrated that our proposed approach consumes less time than a standard routing protocol. Whereas, the latter concentrates on shortest path approach, our model significantly improves time and energy efficiency rates by avoiding link congestions. As would be evident from the figures shown, the initial results are not dissimilar. Both the shortest path method and our proposed approach exhibit identical values in cases where, when links are sufficiently free, our method following W2 factor would also the same route as the shortest path (Figure 3). In other words, when links are free from congestion and adequately capable of handling the route request free from encumbrances, the shortest path approach is also the most efficient at that point of time. It is, however, when the paths are busy, that the W2 factor hands a significant advantage in saving time and minimizing carbon emissions (Figure 4).

Conclusion

In this paper, we have presented our research into on-demand routing based on optimized route selection technique. We have also discussed our proposed algorithm which has been designed to significantly enhance time and energy savings through load balancing while routing mobile network transmissions connected to a cloud environment. Our approach,

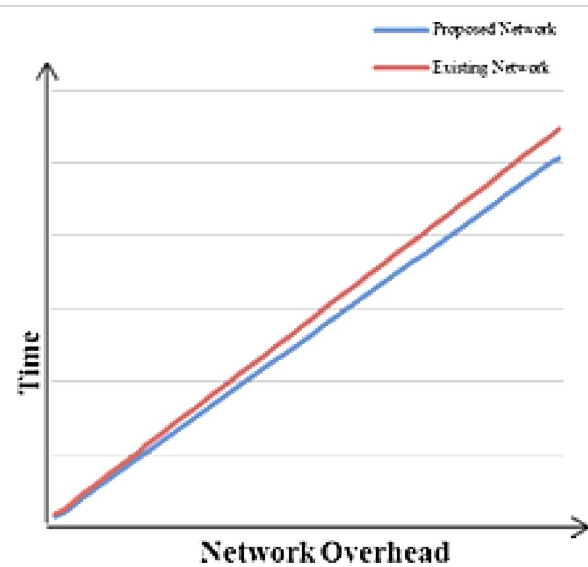


Figure 3 Comparison between existing network and proposed network.

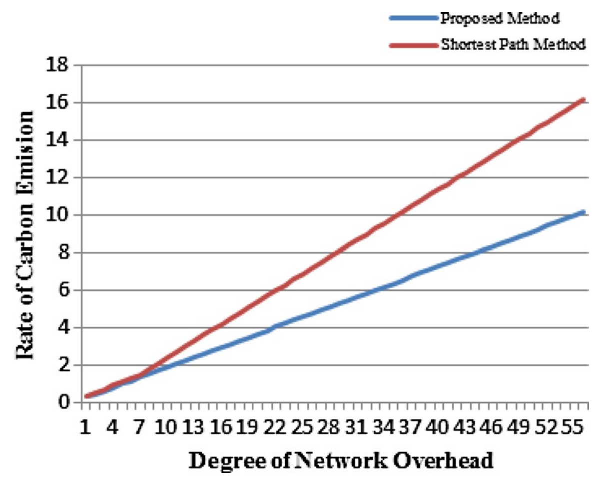


Figure 4 Comparison in terms of carbon emission between existing network and proposed network.

which consists of twin subordinate methods, combined both mobile and cloud networks. Further, our model has been designed to be implemented in an existing mobile cloud computing infrastructure. Our research reveals shortest approach to be more popular with majority authors in our literature survey conducted. However, it suffers a drawback in cases where links are busy. In such situations, the chances of high rates of time and energy consumption beyond acceptable parameters are increased.

The aim of this work is to present a model that is able to address this failing of shortest path routing system. Through our work, we have presented a design to obviate link congestions to the extent possible. In our view, and as can be seen from the analysis of the results obtained, our weight-based base station selection accords a significant advantage in selecting the appropriate route. With the help of our proposed method, idle channels can also be utilized, thereby, evading possible situations involving congested network and waste of energy brought about by resultant carbon emissions.

Using the routing algorithm explained in this paper, data frames that are accumulated can be routed and delivered crisply from cloud data centers to mobile networks or vice versa. This would no doubt be an invaluable tool in enhancing throughout while combining energy efficiency and time-saving features. In future work, we plan to build, design and test more fluid algorithms to handle on-demand routing arising out of multiple devices initiating multiple requests.

Abbreviations

BS: base station; MD: mobile device; MSC: mobile switching center; RM: resource manager; LM: link manager; W1: weight factor regarding base station; W2: weight factor regarding routing.

Authors' contributions

DS provided invaluable assistance in the course of preparing this work. The paper would have been incomplete without his inputs in different section. RB has done background, related work and finally concludes this paper. SS was instrumental in preparing the Algorithm and designing the proposed system. He also prepared the architecture and flowchart which describes the process flow of our proposed design followed by Results and Discussion. All the authors agree that the paper may be submitted in any format appropriate for publishing in the journal. The entire manuscript has been prepared by SS, RB and DS. All authors read and approved the final manuscript.

Acknowledgements
Authors gratefully acknowledge to CSE Department of University of Kalyani and JIS College of Engineering, for providing lab and related facilities for do the research.

Compliance with ethical guidelines

Competing interests
The authors declare that they have no competing interests.

Received: 15 January 2015 Accepted: 17 June 2015
Published online: 27 June 2015

References

Ahmed E, Gani A, Sookhak M, AbHamid SH, Xia F (2015) Application optimization in mobile cloud computing: Motivation, taxonomies, and open challenges. J Network Comp Appl 52:52–68
Cavdar C (2011) Energy-efficient connection provisioning in WDM optical networks. In: Proceedings IEEE Optical Fiber Communication Conference (OFC), Los Angeles, USA, pp 1–5
Cho B, Gupta I (2010) New algorithms for planning bulk transfer via internet and shipping networks. In: Proceedings IEEE ICDCS, 2010
Cho B, Gupta I (2011) Budget-constrained bulk data transfer via internet and shipping networks. In: Proceedings of ACM ICAC, 2011
Chun BG, Ihm S, Maniatis P, Naik M, Patti A (2011) Clone cloud: elastic execution between mobile device and cloud. In: Proceedings of the Sixth Conference on Computer Systems, EuroSys'11, ACM, New York, NY, USA, pp 301–314
Cuervo E, Balasubramanian A, Cho DK, Wolman A, Saroiu S, Chandra R et al (2010) Maui: making smartphones last longer with code offload. In: Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services, MobiSys'10, ACM, New York, NY, USA, pp 49–62
Deboosere L, Simoens P, Wachter JD, Vankeirsbilck B, Turck FD, Dhoedt B et al (2011) Grid design for mobile thin client computing. Future Gener Comp Syst 27:681–693
Dong X, El-Gorashi T, Elmirghani J (2011) IP over WDM networks employing renewable energy sources. J Lightwave Technol 29(1):3–14
Marshall P, Keahey K, Freeman T (2010) Elastic site: using clouds to elastically extend site resources. In: 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing (CCGRID), pp 43–52
Mei L, Chan W, Tse T (2008) A tale of clouds: paradigm comparisons and some thoughts on research issues. In: Proceedings of the Asia-Pacific Services Computing Conference, APSCC'08, IEEE, pp 464–469
Moschakis IA, Karatza HD (2012) Evaluation of gang scheduling performance and cost in a cloud computing system. J Supercomput 59(2):975–992
Park J, Yu H (2012) Resource allocation techniques based on availability and movement reliability for mobile cloud computing. In: Proceedings of the 8th International Conference on Distributed Computing and Internet Technology, ICDCIT'12, Springer-Verlag, Berlin, Heidelberg, pp 263–264
Ricciardi S, Careglio D, Palmieri F, Fiore U, Santos-Boada G, Sol e-Pareta J (2011) Energy-aware RWA for WDM networks with dual power sources. In: Proceedings IEEE International Conference on Communications (ICC 2011), Kyoto, Japan, pp 5–9
Sarddar D, Bose R (2014) A mobile cloud computing architecture with easy resource sharing. Int J Curr Eng Technol 4(3):1249–1254
Satyanarayanan M, Bahl P, Caceres R, Davies N (2009) The case for VM-based cloudlets in mobile computing. IEEE Pervasive Comput 8:14–23
Shen G, Tucker RS (2009) Energy-minimized design for IP over WDM networks. IEEE/OSA J Opt Commun Networking 1(1):176–186