Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IMPACT FACTOR: 5.258

IJCSMC, Vol. 5, Issue. 3, March 2016, pg.120 – 123

JUXTAPOSITION OF METROPOLIS HASTING RANDOM WALK (MHRW) & METROPOLIS HASTING ALGORITHM WITH DELAYED ACCEPTANCE (MHDA)

Bhagyashree.P[1]

[1]M.Tech Student, Department of Computer Science and Engineering
New Horizon College of Engineering
Bangalore, India
[1] bhagyashree.pandian@gmail.com

Abstract--- Online Social Networks (OSNs) have burgeoning growth for the last few years. There are miscellaneous examples of these OSNs. The most prevailing ones are Facebook, Twitter, Flickr, etc. The graphs for symbolizing these immense networks are arduous. Graph sampling is a method which takes a portion (or subset) of the commencing graph and keeps all the features of this commencing graph. There are assorted graph sampling algorithms. In this paper we will focus on MHRW and MHDA graph sampling algorithms and juxtapose their performance.

Keywords--- OSNs, Graph sampling, MHRW, MHDA, NMSE

I. INTRODUCTION

Online Social Networks (OSNs) are in-demand crux for intimation and propaganda sharing over the Internet. It provides interconnection among diverse users. These OSNs contribute manifesto for administering events, set up social links to soul mates, to share resources or materials among your playmates etc. These OSNs have gained dramatic contemplation from every nook of the world. Examples: Facebook, Twitter, Flickr, YouTube, LiveJournal, MySpace, Orkut, Linkedln etc. The predominant perception of social search is to manoeuvre facts serene from a users' social network to enhance the meticulousness of search results. Social search has acquired vigilance as a method towards personalized search.

Nevertheless, for stupendous Online Social Networks, typically comprising millions of users, a detailed creep of user's amplified neighbourhood is impractical. Accordingly, proficient or competent procedures are required. One such technique that was introduced is graph sampling. Graph sampling is a method which takes a portion of the commencing graph and keeps all the features of this commencing graph. In accordance with this technique, we choose compact or small sample of the commencing graph and examine it. This technique that is graph sampling technique provides an effective and modest solution.

Graph sampling technique has a dilated spectrum of applications. Example: survey hidden population in sociology, visualize social graph, scale down Internet as graph, graph sparsification, etc. There are diversified graph sampling algorithms. They are Vertex sampling, Edge sampling, Traversal Based sampling, Breadth First Search (BFS) sampling, Frontier sampling (FS), Random Walk (RW), Metropolis Hasting Random Walk

(MHRW), Re-Weighted Random walk (RWRW), Metropolis Hasting with Delayed Acceptance (MHDA), Forest Fire (FF), Snowball Sampling (SBS), etc.

In this paper we will focus on Metropolis Hasting Random Walk (MHRW) and Metropolis Hasting algorithm with Delayed Acceptance (MHDA) graph sampling algorithms and juxtapose their performance.

II. METROPOLIS HASTING RANDOM WALK (MHRW)

Metropolis Hasting Random Walk objective is to probe closest successive neighbour. For example: The figure1 is as shown below: Say you're traversing through the graph from source node as India and destination node as Pakistan. This is the general flow of MHRW algorithm. There are four possible avenues to reach the destination from the source as you can see in the figure. The first way is via Yemen. The second way is direct to the destination. The third way is via Dubai. And the fourth way is via Nepal. According to MHRW algorithm, the source node (India) will randomly opt the closest successive neighbour. So if the journey via Yemen is closer than other three routes, then the algorithm chooses the Yemen as the next node to be visited to reach the destination. The idea that is conveyed from this algorithm is the node that is closer will be selected to be visited. This concept is carried over throughout till the destination is reached.

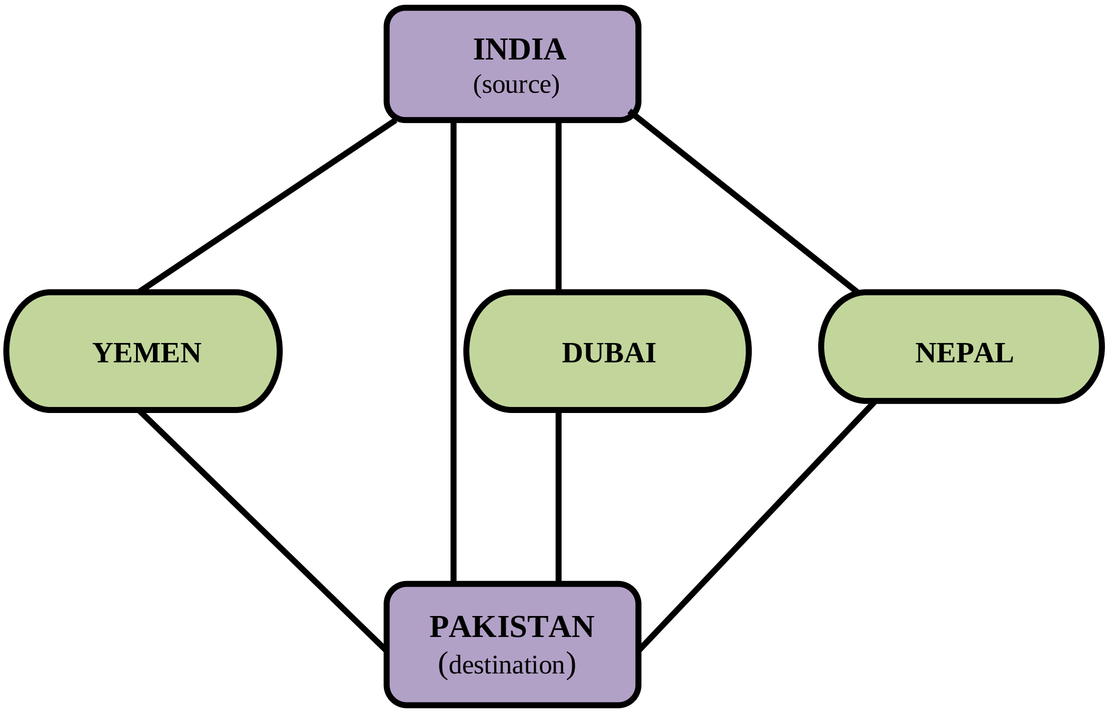

Fig.1. Demonstration of Algorithms

MHRW algorithm is explained in detail with the help of equations. Let f(x) be the target density, $x^{(j)}$ be the current value and $q(x|x^{(j)})$ be the proposal distribution that depends on the current value $x^{(j)}$. We have the sample $x^* \sim x^{(j)}$ where x^* is the proposed value and $q(x|x^{(j)})$ is the proposal distribution. Then we calculate the acceptance probability using the formula $p(x^{(j)}, x^*) = \min\left\{1, \frac{f(x^*)}{f(x^{(j)})}\frac{q(x^{(j)}|x^*)}{q(x^*|x^{(j)})}\right\}$. Where $\frac{f(x^*)}{f(x^{(j)})}$ the target density of proposed value (numerator) is verses the current value (denominator). $\frac{q(x^{(j)}|x^*)}{q(x^*|x^{(j)})}$ is the proposal density value. The denominator of this equation is how likely the proposed value was given on the current state. The numerator of this equation is how likely the current value would occur if we are in the proposed state. If the calculated acceptance probability is successful then we set $x^{(j+1)} = x^*$ with the probability $p(x^{(j)}, x^*)$. Otherwise set $x^{(j+1)} = x^{(j)}$.

Now observe the following equation of MHRW which was explained earlier. $p(x^{(j)}, x^*) = \min\{1, \frac{f(x^*)}{f(x^{(j)}}\frac{q(x^{(j)}|x^*)}{q(x^*|x^{(j)})}\}$. The second part of the above equation that is, $\frac{q(x^{(j)}|x^*)}{q(x^*|x^{(j)})}$ is symmetric. So it cancels each other. Then we obtain the following equation. $p(x^{(j)}, x^*) = \min\{1, \frac{f(x^*)}{f(x^{(j)})}\}$. If $f(x^*)$, probability value is higher than $f(x^{(j)})$ then it is always accepted. Otherwise it is accepted with some probability.

III. METROPOLIS HASTING ALGORITHM WITH DELAYED ACCEPTANCE (MHDA)

Metropolis Hasting Algorithm with Delayed Acceptance (MHDA) is hypothetically promised to accomplish no additional cost for biased and unbiased graph sampling. It also contributes to higher efficiency. A significant factor of the MHDA algorithm is its suitability for any non-uniform node sampling. It certifies superior sampling capability. An extraordinary characteristic of MHDA algorithm is its universal objective. This algorithm recollects the formally visited node from which the random walk came. MHDA algorithm is explained with the figure above. This algorithm works similar to MHRW, but it keeps the track of the nodes that are visited. So it helps to backtrack also.

Suppose $q(x|x^{(j)}) = q^{(x)}$, then we call q independent. The acceptance probability is given by the following formula: $p(x^{(j)}, x^*) = \min\{ 1, \frac{f(x^*)}{f(x^j)} \frac{q(x^j|x^*)}{q(x^*|x^j)} \}$. In this equation, the term $\frac{q(x^j|x^*)}{q(x^*|x^j)}$ does not depend on any proposal value because the proposal value is independent.

IV. JUXTAPOSITION OF MHRW & MHDA

Juxtaposition of MHRW and MHDA is shown with a graphical simulation in the below figure 2. The input considered is Wikipedia dataset. MHDA contributes no additional cost as MHRW. The competence and the productivity are eminent when we deal with MHDA. Based on the factors like cost, area, axis, efficiency, distance, and NMSE, MHDA shows a worthier performance when juxtaposed to MHRW.

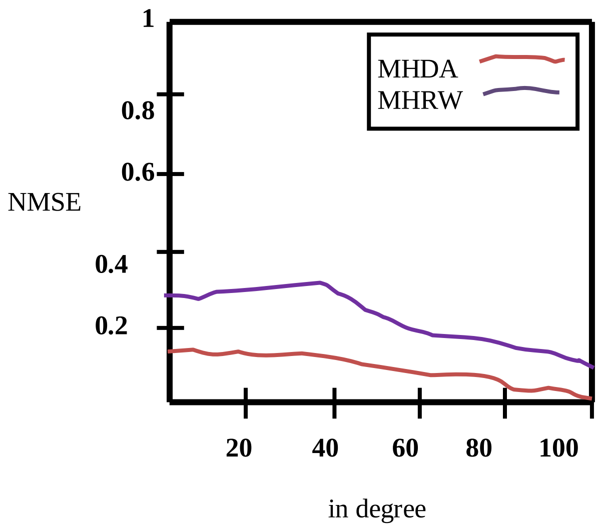

Fig.2. Juxtaposing of MHDA & MHRW in terms of NMSE

Normalised Mean Square Error (NMSE) is the error that shows immense performance when error is low. As observed in the figure 2, MHDA shows immense performance than MHRW since it has low error than MHRW.

V. CONCLUSION

The objective of this project is to show the juxtaposition of Metropolis Hasting Random Walk (MHRW) and Metropolis Hasting Algorithm with Delayed Acceptance (MHDA). This juxtaposition is shown in terms of Normalized Mean Square Error (NMSE). With help of simulation of graph (in figure 2), the juxtaposing of these two algorithms is clearly differentiated. As observed in the graph (in figure 2), it is concluded that MHDA performs better than MHRW algorithm.

REFERENCES

[1] R. Rajkumar, " Behaviour Analysis and feature selection in Online Social Network", In Proc. Of ACM, 2015.
[2] Marco Banterle," Accelerating Metropolis Hastings alogorithms: Delayed acceptance with prefetching", arXiv: 1406.2660v[stat,CO]10 Jun 2014.
[3] Korattikara, Chen, and Welling (2013), Austeririty in MCMC land: " Cutting the Metropolic-Hastings budget", arXiv:1304.5299.
[4] M. Kurant, M. Gjoka, Y. Wang, Z. Almquist, C. Butts, and A. Markoloulou, " Coarse-grained topology estimation via graph sampling", IN Workshop on OSNs, Helsinki, Norway, August 2012.

[5] A. S. Maiya and T. Y. Berger-Wolf, "Benefits of Bias: Towards Better Characterization of Networking Sampling". In Proc. of KDD, San Diego, CA,USA, August 2011.

[6] M. Gjoka, "Practical Recommendations on Crawling Online Social Networks", vol. 29, NO, 9, October 2011.

[7] M.Gjoka M Kurant, C.T.Butts, and A Markopoulou, "Walking in Facebook: A Case Study of Unbiased Sampling of OSNsm"In Proc. of IEEE INFOCOM, 2010.

[8] Strid, "Efficient parallelization of Metropolis Hasting Algorithms using a Pre-fetching Approach", Computational Statistics & Data Analysis, 2010.

[9] P.Berenbrink, C.Cooper, T.R.R.Els asser and T.Sauerwald, "Speeding up random walks with neighbourhood exploration". In ACM SODA, 2010.

[10] M. A. Hasan and M.J.Zaki, "Output space sampling for graph patterns". In VLDB, 2009.

[11] J. Leskovec and C.Faloutsos, "Sampling From Large Graphs", In KDD, 2006.

[12] V.Krishnamurthy, M.Faloustsos, M.Chrobak, L.Lao, J.Cui, and A.Percus, "Reducing large internet topologies for faster simulations". In Proceedings of the IFIP Networking, Waterloo, Canada, 2005.

[13] S. Itzkovitz, R. Milo, N. Kashtan, G. Ziv, and U. Alon, "Sub-graphs in random networks", Physical Rev. E, 68, 2003.

[14] P.J.Green and A.Mira, "Delayed rejection in reversible jump metropolis-hastings", Biometrika, 88(4): 1035-1053, 2001.

[15] A.Mira. "Ordering and improving the performance of monte carlo markov chains". Statistical Science, 16(4):340-350,2001.

[16] Christian H"ubler, Hans-Peter Kriegel, "Metropolis Algorithms for Representative Subgraph Sampling".

[17] Chul-Ho Lee,"Beyond Random Walk and Metropolis-Hastings Samplers".

