American Journal of Engineering Research (AJER)
e-ISSN : 2320-0847 p-ISSN : 2320-0936
Volume-03, Issue-04, pp-281-294
www.ajer.org

Research Paper Open Access

Probabilistic Stochastic Graphical Models With Improved Techniques

[1]Swati Kumari, [2]Anupama Verma
Assistant Professor, JRU, Department of Computer Science Engineering, Ranchi - 834002
Assistant Professor, JRU, Department of Computer Science Engineering, Ranchi – 834002

Abstract: - The web may be viewed as a directed graph. A HTML page can be treated as node and hyperlink as an edge from one node to another this will form a directed graph. But the nature of this graph is evolving with time, so we require evolving graph model. The $G_{n,p}$ and $G_{n,N}$ are static models. In these models, graph G = (V,E) is not changing with time. Sometimes we need a graph model which is time evolving Gt = (V_t,E_t). There are two characteristic functions of evolving graph model [5] first gives the number of vertex added at time t+1, whereas second gives set of edges added at time t+1. Evolving copying model is different from traditional $G_{n,p}$ in the following:
1. Independently chosen edges do not show the results found on the web.
2. This model captures the evolving natures of web graph.
In this paper we review some recent work on generalizations of the random graph aimed at correcting these shortcomings. We describe generalized random graph models of both directed and undirected networks that incorporate arbitrary non-Poisson degree distributions, and extensions of these models that incorporate clustering too. We also describe two recent applications of random
graph models to the problems of network robustness and of epidemics spreading on contact networks. This paper studies the random graph models. Concentration is made over the paper stochastic models for web graph by Ravi Kumar et. al. [5]. During the proof of the results given in the paper some more precise results have been found. Some of results are here which can be compared with results of the papers. The improvements are presented here along with results in the original paper for comparison.

Keywords: *- Random Graph, Martingale, Web Graph.*

I. INTRODUCTION

Definition 1. *A HTML page can be treated as node and hyperlink as an edge from one node to another this will form a directed graph called web graph*.

At present this web graph has billion vertices [5], and an average degree of about 7. In this thesis we discuss different random graph models. These observation suggest that web is not well modelled by traditional models such as *Gn,p*. A lot of models for the random graph have been proposed like *Gn,p* and *Gn,N* models [6] etc. These models do not ensure the power-law for degree of vertices and do not explain the abundance of bipartite cliques observed in the web graph. The copying graph model [5] explains these and also covers the evolving nature of the web graph.
Evolving copying model is different from traditional Gn,pin the following:
(a) Independently chosen edges do not show the results found on the web.
(b) This model captures the evolving natures of web graph.
In this model during every time interval one vertex arrives. At time step t, a single vertex u arrives and the i-th out-link of u is then chosen as follows. With probability α, the destination is chosen uniformly at random from V_t, with remaining probability the outlink is taken to be the i^{th} outlink of prototype vertex p. where V_t is set of vertex at time t.

A. Problem Definition

The problem is to develop random graph model that explain the different nature of web graph. We have to study probabilistic techniques, different models for random graph and find more results based on the models

B. Related Work

The "copying" model analysed in this thesis were first introduced by Kleinberget. al. [4]. Motivated by observations of power-laws for degrees on the graph
of telephone calls, Aiello, Chung, and Lu [1] propose a model for "massive graph"(Henceforth the "ACL model"). In 2000, Ravi Kumar et. al. proposed a numberof models in his paper "Stochastic models for the web graph" [5]. This thesis hassome more precise results, and also the proof of different results available in thepaper "Stochastic models for the web graph".

C. Contribution

(1) The web graph is so large that the study of different nature of this graph is difficult without a good modeling.
(2) The evolution of different social networking site can be explained with the help of this modeling.

D. Results and Organization

Results in Paper	Mywork done
$\lvert E[N_{t,0} \lvert N_{t-k,0}] - E[N_{t,0} \lvert N_{t-(k+1),0}]\rvert \leq 2$	$\lvert E[N_{t,0} \lvert N_{t-k,0}] - E[N_{t,0} \lvert N_{t-(k+1),0}]\rvert \leq 1$
$\frac{t+\alpha}{1+\alpha} - \alpha^2 - \ln t \leq E[N_{t,0}] \leq \frac{t+\alpha}{1+\alpha}$	$E[N_{t,0}] = \frac{t}{1+\alpha}$
$Pr\lvert N_{t,0} - E[N_{t,0}] \geq l\rvert \leq \exp(-\frac{l*l}{4t})$	$Pr\lvert N_{t,0} - E[N_{t,0}] \geq l\rvert \leq \exp(-\frac{l*l}{2t})$
	$E[N_{t,0}] = \sum_{j=0}^{t-1} Sj, 0 = \frac{t}{1+\alpha} \frac{t}{1+\alpha}$
	$P_t = \lim_{t\to\infty} \frac{E[Nt,0]}{t} = \frac{1}{1+\alpha}$
$\lvert E[N_{t,i} \lvert N_{t-k,i} N_{t-k,i-1}] - \lvert E[N_{t,i} \lvert N_{t-(k-1),i} N_{t-(k-1),i-1}]\rvert \leq 2$	$\lvert E[N_{t,i} \lvert N_{t-k,i} N_{t-k,i-1}] - \lvert E[N_{t,i} \lvert N_{t-(k-1),i} N_{t-(k-1),i-1}]\rvert \leq 1$ for $N_{t-k,I} = N_{t-(k+1),i}$

II. PROBABILISTIC TECHNIQUES

Preliminaries

We define a discrete probability space (Ω, p) where Ω is discrete (finite or countable infinite) set and p is probability such that
$P: \Omega \to [0,1]$means$\forall\ \omega \in \Omega\ \ p(\omega) \in [0, 1]$ and

$$\sum_{\omega} p(\omega) = 1$$

In this thesis we will discuss only discrete probability spaces.
Definition 2. *Given $(\Omega, p), B \subseteq \Omega$, Define conditional probability over $(\Omega, p(B))$as follows.*

$$p(\omega/B) = \begin{cases} 0\ if\ \omega \notin B, \\ \frac{p(\omega)}{p(B)}\ if\ \omega \epsilon B \end{cases}$$

Where $p(\omega/B)$ is the probability of happening of ω when B has already happened?
Now the probability of happening of A when B has already happened is denoted by P(A|B) and is given by

$$P(A|B) = \sum_{\omega \epsilon A} p(\omega|B)$$

$$= \sum_{\omega \epsilon A \cap B^C} p(\omega|B) + \sum_{\omega \epsilon A \cap B} p(\omega|B)$$

$$= 0 + \sum_{\omega \epsilon A \cap B} p(\omega|B)$$

$$= \sum_{\omega \epsilon A \cap B} \frac{p(\omega)}{P(B)}$$

$$= \frac{1}{p(B)} \sum_{\omega \epsilon A \cap B} p(\omega)$$

$$\Rightarrow P(A|B) = \frac{p(A \cap B)}{p(B)}$$

Expectation of Random Variable

Expectation of a random variable is its basic characteristic. The expectationof random variable is weighted average of the values it assumes, where each valueis weighted by the probability that the variable assumes that value.

Definition 3. *A random variable X is a real-valued function over the sample space Ω; that is X :Ω → R. The expectation of random variable X is defined as*

$$E\{X\} = \sum_j X(\omega_j)p(\omega_j)$$

The expectation is finite If $\sum_j |X_{\omega_j}|p(\omega_j)$converges; otherwise, the expectationis unbounded.

For example:A random variable X that takes the values 2^j with probability $\frac{1}{2^j}$ for j = 1,2,……

The expectation of X

$$E[X] = \sum_{j=1}^{\infty} \frac{1}{2^j} 2^j \rightarrow \infty$$

Here the expectation is unbounded.

Conditional Expectation

Definition 4.*The conditional expectation [3] of X relative to B is defined whenP{B} > 0 as*

$$E(X|B\} = \sum_j X(\omega_j)p(\omega_j|B)$$

Definition 5.*Let (Ω_1; p1) and (Ω_2; p2) be probability distributions and (Ω; p) iscalled the joint distribution of them if the following conditions are satisfied.*

$$\sum_{(\omega_1,\omega_2)\epsilon\Omega} p(\omega_1,\omega_2) = 1$$

$$\sum_{\omega_1} p(\omega_1,\omega_2) = P_2\ \omega_2 = p(\omega_2)$$

$$\sum_{\omega_2} p(\omega_1,\omega_2) = P_1\ \omega_1 = p(\omega_1)$$

where

$$\Omega = \Omega_1 \times \Omega_2 = \{(\omega_1,\omega_2)|\omega_1\epsilon\Omega_1,\omega_2\epsilon\Omega_2\}$$

Definition 6.Two *random variable X and Y are independent if and only if*

$$Pr\big((X = x) \cap (Y = y)\big) = Pr(X = x) . Pr(Y = y)$$

for a values of x and y. Similarly, random variables X_1,X_2,…,X_k are mutuallyindependent if and only if, for any subset I ⊆ [1, k] and any values x_i, i∊I,

$$Pr\left(\bigcap_{i\epsilon I} X_i = x_i\right) = \prod_{i\epsilon I} Pr(X_i = x_i)$$

If the events ξ_i and ξ_j are pairwise independent then

$$Pr[\xi_i|\xi_j] = Pr[\xi_i]$$

for all i≠j

Proof.

$$Pr[\xi_i|\xi_j] = \frac{Pr[\xi_i \cap \xi_j]}{Pr[\xi_j]}$$

$$= \frac{\Pr[\xi_i|\Pr[\xi_j]}{\Pr[\xi_j]} = \Pr[\xi_i]$$

Linearity of Expectation

$$E[X + Y] = E[X] + E[Y]$$

Proof.

$$E[X+Y] = \sum_{(x,y)\epsilon\mathbb{R}\times\mathbb{R}} (x+y)p(X,Y)$$
$$= \sum_x \sum_y (x+y)\Pr((X=x)\cap(Y=y))$$
$$= \sum_x \sum_y xPr\big((X=x)\cap(Y=y)\big) + \sum_x \sum_y yPr\big((X=x)\cap(Y=y)\big)$$
$$= \sum_x xPr(X=x) + \sum_y yPr(Y=y)$$
$$= E[X] + E[Y]$$

General form

$$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} [X_i]$$
$$E[aX+b] = aE[X]+b$$

Proof.

$$E[aX+b] = \sum_{x\epsilon\mathbb{R}} (a(X=x)+b)\Pr(X=x)$$
$$= a\sum_{x\epsilon\mathbb{R}} (X=x)\Pr(X=x) + b\sum_{x\epsilon\mathbb{R}} \Pr(X=x)$$
$$= aE[X] + b.1$$
$$= aE[X] + b$$

If X and Y are two independent random variables, thenE[X,Y] = E[X].E[Y]

Bernoulli Random Variable

If we run an experiment such that it succeedswith probability p and fails with probability (1-p). Let Y be a randomvariable such that

$$Y = \begin{cases} 1\ if\ success\ wp\ p, \\ 0\ if\ fails\ wp\ (1-p) \end{cases}$$

Here the random variable Y is called a Bernoulli random variable.

Theorem 1. *The expectation of Bernoulli random variable is same as the probability of success of the random variable.*

Proof.

E[Y] = 1. p + 0. (1 - p) = p

$\Rightarrow$ *E[Y] =Pr[Y=1]*

Binomial Distribution Consider a sequence of n independent experiments,each of which succeeds with probability p. If we represents X as numberof successes in n experiments then X has a binomial distribution.

Definition 7. *A binomial random variable X with parameters n and* p, *denotedby* $B(n,p)$*is defined by following probability distribution on* $j = 0,1,2,\ldots,n$:

$$Pr(X=j) = \begin{cases} \binom{n}{j} p^j (1-p)^{n-j} & k\epsilon\{0,1,2,\ldots.n\} \\ 0\ otherwise \end{cases}$$

Theorem 2.*The expectation of binomial random variable is* np.

Proof. The binomial random variable X can be expressed as the sum of Bernoulli random variables.

$$X = \sum_{i=1}^{n} X_i$$
$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$$

from Linearity of expectation

$$\sum_{i=1}^{n} p$$

from expectation of Bernoulli random variable

$$= np$$

Geometric Distribution

Sequence of independent trials until the first success is found where each trialsuccess with probability p, This gives an example of geometric distribution.

A geometric random variable X with parameter p is given by following probabilitydistribution on n = 0,1,2,….

$$\Pr(X = n) = (1 - p)^{(n-1)}.p$$

For geometric random variable X equals n there must be $n - 1$ failures followedby a success.

Definition 8. $\sum_{i=1}^{n} \frac{1}{t} = \mathbb{H}(n)$*called the harmonic number* $= \ln n + \Theta(1)$.

Proof. Since $\frac{1}{i}$ is monotonically decreasing function, we can write $\ln n = \int_{x=1}^{n} \frac{1}{k} \leq \sum_{k=1}^{n} \frac{1}{k}$

$$\Rightarrow \sum_{k=2}^{n} \frac{1}{k} \leq \int_{x=1}^{n} \frac{1}{k} = \ln n$$
$$\Rightarrow \ln n \leq \mathbb{H}(n) \leq \ln n + 1$$
$$\Rightarrow \mathbb{H}(n) = n \ln n + \Theta(n)$$

Theorem 3.*The expectation of geometric random variable is* $\frac{1}{p}$

Proof.

$$E[X] = \sum_{i=1}^{\infty} i.(1-p)^{i-1}.p$$
$$= p.\sum_{i=1}^{\infty} i.(1-p)^{i-1}$$
$$= p.\sum_{i=1}^{\infty} i.t^{i-1} \quad (\because 1 - p = t\ 0 < t < 1)$$
$$= p.(1 + 2t + 3t^2 + 4t^3 + \cdots)$$
$$= p.\frac{1}{(1-t)^2} = p\frac{1}{p^2} = \frac{1}{p}$$
$$\{\because f(x) = 1 + 2t + 3t^2 + 4t^3 + \cdots = \frac{1}{1-t} \quad sum\ of\ G.P.$$
$$f'(x) = 1 + 2t + 3t^2 + 4t^3 + \cdots = \frac{1}{(1-t)^2}\}$$

Markov's Inequality

Let X be a random variable assumes only non-negative values the for all $t > 0$

$$\Pr[X \geq t] \leq \frac{E[X]}{t}$$

Proof. $E[X] = \sum_{x \epsilon \mathbb{R}} xPr(X = x)$

$$\geq \sum_{x \geq k} xPr(X = x)$$
$$\geq \sum_{x \geq k} kPr(X = x)$$
$$E[X] \geq \sum_{x \geq k} kPr(X = x) = kPr(X \geq k)$$
$$E[X] \geq kPr(X \geq k)$$
$$Pr(X \geq k) \leq \frac{E[X]}{k}$$
$$\Longrightarrow Pr(X \geq t) \leq \frac{E[X]}{t}$$

Chebyshev's Inequality

$$\Pr(|X - E[X]| \geq t) \leq \frac{Var[X]}{t^2}$$

Proof.

$$\Pr(|X - E[X]| \geq t) = \Pr((|X - E[X])^2| \geq t^2)$$
$$\leq \frac{(X - E[X])^2}{t^2}$$
$$= \frac{Var[X]}{t^2}$$

Since $(X - E[X])^2$ is a non-negative random variable, we can apply Markov'sinequality.

Jensen's Inequality

If f is a convex function, then $E[f(X)] \geq f(E[X])$|

Proof. Suppose f has a Taylor series expansion. Expending $\mu = E[X]$and usingthe Taylor series expansion with a remainder term, yields for some a.

$$f(x) = f(\mu) + f'(\mu)(x-\mu) + \frac{f''(a)(x-\mu)^2}{2}$$
$$\geq f(\mu) + f'(\mu)(x-\mu) \ \ (\because f''(a) \geq 0 \ by\ convexity.)$$
$$f(x) \geq f(\mu) + f'(\mu)(x-\mu)$$

Taking expectation on both sides

$$E[f(x)] \geq f(\mu) + f'(\mu)E[(x-\mu)] = f(\mu)$$
$$E[f(x)] \geq f(E[X])$$

Martingale

σ- field

Definition 9.*A σ- field$(\Omega, \mathbb{F})$ consists of a sample space and a collection ofsubsets $\mathbb{F}$ satisfying the following conditions [7].*

$$\phi \epsilon \mathbb{F}$$
$$\varepsilon \epsilon \mathbb{F} \Rightarrow \varepsilon' \epsilon \mathbb{F}$$
$$\varepsilon_1, \varepsilon_2, \dots \epsilon \mathbb{F} \Rightarrow \varepsilon_1 \cup \varepsilon_2 \cup \dots \epsilon \mathbb{F}$$

Partition of Ω

Definition 10.$\mathbb{F}$is a partition of Ω if $\mathbb{F} \subseteq 2^{\Omega}$ and

(i) $\bigcup_{F \epsilon \mathbb{F}} F = \Omega$

(ii) $F \neq \phi \ \forall \ F \epsilon \mathbb{F}$

(iii) $F \cap G = \phi \ \forall F \neq G \ where\ F, G \epsilon \mathbb{F}$

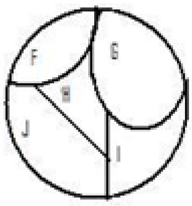

Figure 1: Partition of Sample space

Here $\mathbb{F}$ is partition of Ω then corresponding algebra of partition $\mathbb{A}(\mathbb{F}_1)$ is

$$\{\phi, F, G, H, \dots, F \cup G, F \cup H, \dots, F \cup G \cup H, \dots\}$$

Where algebra is smallest set containing $\mathbb{F}$ that is closed under finite union, intersectionand complementation.

Definition 11. *If $\mathbb{F}_1, \mathbb{F}_2$ are partitions of Ω then $\mathbb{F}_1 \subseteq \mathbb{F}_2$ if $A \epsilon \mathbb{A}(\mathbb{F}_1)$implies $A \epsilon \mathbb{A}(\mathbb{F}_2)$. Where $\mathbb{A}(\mathbb{F}_1)$ and $\mathbb{A}(\mathbb{F}_2)$ are algebra of $\mathbb{F}_1$and $\mathbb{F}_2$ respectively.*

Filtration

Definition 12*. Given the σ- field $(\Omega, \mathbb{F})$ with $\mathbb{F} \subseteq 2^{\Omega}$, a filter (sometimes alsocalled filtration) is a nested sequence $\mathbb{F}_1 \subseteq \mathbb{F}_2 \subseteq \mathbb{F}_3 \dots \subseteq \mathbb{F}_n$ of subsets of 2^{Ω}suchthat [7]* $\mathbb{F}_0 = \{\phi, \Omega\}$

$$\mathbb{F}_n = 2^{\Omega}$$
$$for\ 0 \leq i \leq n, (\Omega, \mathbb{F}_i)\ is\ a\ \sigma -\ field.$$

Definition 13. Let $(\Omega, \mathbb{F})$ be any $\sigma - field$, and Y any random variable that takeson distinct values on the elementary events in $\mathbb{F}$. Then $E[X|\mathbb{F}] = E[X|Y]$.

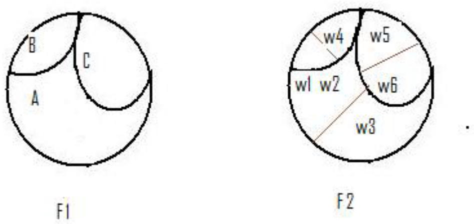

Figure 2: Filtration

Notice that the conditional expectation $E[X|Y]$does not really depend on theprecise value of Y on the specific elementary event. In fact, Y is merely an indicatorof the elementary events in$\mathbb{F}$. Conversely, we can write

$E[X|Y] = E[X|\sigma(Y)]$ where $\sigma(Y)$ is the $\sigma-$ $field$ generated by the events of the type$\{Y = y\}$, i.e., thesmallest σ- field over which Y is measurable.
Definition 14. *A random variable X over σ- field* $(\Omega, \mathbb{F})$ *is well defined if*$X(\omega_1) = X(\omega_2)$*for all* $A\epsilon\mathbb{F}$ *where* $\omega_1, \omega_2 \epsilon A$ *and* $\mathbb{F}$ *is partition.*
It is also called $\mathbb{F}$ *-measurable. A random variable X is* $\mathbb{F}$ *-measurable if its valueis constant over each block in the partition generation* $\mathbb{F}$*. Since the partitions generating the σ- fields in a filter are successively more refined, it follows that if* $X, \mathbb{F}_i$ *-measurable, it is also* $\mathbb{F}_j$*-measurable for all* $j \geq i$*. i.e*
If $\mathbb{F}_1 \subseteq \mathbb{F}_2$ *and X is well defined over* $(\Omega, \mathbb{F}_1)$ *then X is well defined over*$(\Omega, \mathbb{F}_2)$ *Converse may not be true.*
Suppose X is well defined over $\mathbb{F}_2$. How to define a random variable over $\mathbb{F}_1$so that it is well defined.

Let random variable

$$Y(\omega) = \frac{\sum_{\omega\epsilon A} X(\omega)p(\omega)}{p(A)} \text{ where } \omega\epsilon A\epsilon\mathbb{F}_1.$$

$$\Rightarrow Y(\omega) = \sum_{\omega\epsilon A} X(\omega)p(\omega|A)$$

Proof. $Y(\omega_i) = \sum_{\omega_i\epsilon A} X_i(\omega)p(\omega_i|A)$

$$= \frac{\sum_{\{i|\omega_i\epsilon A\}} x_i p_i}{\sum_{\{i|\omega_i\epsilon A\}} p_i}$$

$Y(\omega_i)$is same for a partition. So random variable is constant for a particularpartition. Similarly we can proof for other partitions. Here we assume that the Sample space has been partitioned, the points $\omega_1, \omega_2, \omega_3$with probability $p_1, p_2, p_3$and valuesx_1, x_2, x_3 respectively are in finer partition $\mathbb{F}_2$
Theorem 4. Defining Y as conditional expectation of X given $\mathbb{F}_1$ denoted by$E[X|\mathbb{F}_1] = Y$ then

$$E[Y] = E[X]$$

Proof. $E[Y] = \sum_{\omega\epsilon\Omega} Y(\omega)p(\omega)$

$$= \sum_{A\epsilon\mathbb{F}_1}\sum_{\omega\epsilon A} Y(\omega)p(\omega)$$

$$= \sum_{A\epsilon\mathbb{F}_1} Y(\omega)\sum_{\omega\epsilon A} p(\omega) \; As\ Y\ is\ well\ defined$$

$$\sum_{A\epsilon\mathbb{F}_1} Y(\omega)p(A)$$

$$= \sum_{A\epsilon\mathbb{F}_1} p(A)\sum_{\omega\epsilon A} X(\omega)p(\omega|A)$$

$$= \sum_{A\epsilon\mathbb{F}_1} p(A).\frac{\sum_{\omega\epsilon A} X(\omega)p(\omega)}{p(A)}$$

$$= \sum_{A\epsilon\mathbb{F}_1}\sum_{\omega\epsilon A} X(\omega)p(\omega)$$

$$= \sum_{\omega\epsilon\Omega} X(\omega)p(\omega)$$

$$= E[X]$$

From this result it is clear that for filter $\mathbb{F}_1 \subseteq \mathbb{F}_2 \subseteq \mathbb{F}_3 \ldots \subseteq \mathbb{F}_n$ over*σ- field* and corresponding random variables are $X_0, X_1, X_2, \ldots, X_n$then

$$\mu = E[X_0] = E[X_1] = E[X_2] = \cdots = E[X_n]$$

Martingale
Definition 15.*Let* $(\Omega, \mathbb{F}, Pr)$ *be a probability space with a fiter*$\mathbb{F}_0, \mathbb{F}_1, \mathbb{F}_2, \ldots$*.Suppose* $X_0, X_1, X_2, \ldots$ *are random variables such that for all* $i \geq 0$*, Xi is* $\mathbb{F}_i$*-measurable. The sequence* $X_0, X_1, X_2, \ldots, X_n$*is a martingale provided, for all* $i \geq 0$,

$$E[X_{i+1}|\mathbb{F}_i] = X_i$$

Definition 16.*The set of random variables* $\{Z_n, n = 0,1,2 \ldots\}$ *is said to be amartingale [8] with respect to sequence* $\{X_n, n = 0,1,2 \ldots\}$ *if* Z_n *is a function of*$X_0, X_1, X_2, \ldots, X_n E[|Zn|] < \infty$*,and*$E[Z_{n+1}|X_0, X_1, X_2, \ldots, X_n] = Z_n$
If we say $\{Z_n\}$ *is a martingale [8] (without specifying* $\{X_n, n = 0,1,2 \ldots\}$ *whenit is a martingale with respect to itself. i.e.* $\{Z_n\}$ *is a martingale if* $E[|Zn|] < \infty$*,and*$E[Z_{n+1}|Z_0, Z_1, Z_2, \ldots Z_n] = Z_n$.
Note:-$\{Z_n\}$ is martingale with respect to $\{X_n\}$ then it is a martingale.
Theorem 5.$\{Z_n\}$is martingale with respect to $\{X_n\}$ then it is a martingale

$$E[Z_{n+1}|Z_0,Z_1,Z_2,\dots Z_n] = Z_n$$

*Proof.*we know the identity $E[X|U] = E[E[X|U,V]|U]$

$\{Z_n\}$is martingale with respect to $\{X_n\}$ then

$$\begin{aligned} &E[Z_{n+1}|Z_0\dots Z_n] \\ &= E[E[Z_{n+1}|Z_0\dots Z_n,X_0\dots X_n]|Z_0\dots Z_n] \\ &= E[E[Z_{n+1}|X_0\dots X_n]|Z_0\dots Z_n] \\ &= E[Z_n|Z_0\dots Z_n] \\ &= Z_n \end{aligned}$$

Theorem 6. $E[Z_n] = E[Z_0]$

Proof. $Z_0\dots Z_n$are functions of $X_0\dots X_n$

$$E[Z_{n+1}|Z_0,Z_1,Z_2,\dots Z_n] = Z_n$$

Taking expectation on both the sides.

$$E[Z_{n+1}] = E[Z_n]$$
$$E[Z_n] = E[Z_0]$$

$E[Z_0]$is called mean of the martingale.

Example: - Fair bet game:suppose X_i be the outcome of i^{th} game. $Z_n$is the fortune of gambler after n^{th}game. For given outcome of first n games the gambler's expected fortune after(n + 1)st game is equal to his fortune before the game. The gambler's expectedwinning on each game is equal to zero. Here $X_0,X_1,X_2,\dots,X_n$ is martingale.

Azuma's Inequality

Let $X_0,X_1,X_2,\dots,X_n$ be a martingale sequence such that for all k

$$|X_k - X_{k-1}| \le c$$

$$\Pr[|X_t - X_0| \ge l] \le 2\exp(-\frac{l^2}{2tc^2})$$

Where c may depend on k for all t≥0 and for any l>0

Proof. Let $\mathbb{F}_1 \subseteq \mathbb{F}_2 \subseteq \mathbb{F}_3 \dots \subseteq \mathbb{F}_n$ be filter sequence and $X_0,X_1,X_2,\dots,X_n$ be martingalesequence where $X_0 = E[X_i]$ for $i\epsilon 0,1,2,\dots,n$. Also

$$E[X_j|\mathbb{F}_i] = X_i \quad for\ i < j$$

If X is $\mathbb{F}_i$ measurable (constant for particular partition) and Y is anotherrandom variable then we observe that

$$E[XY|\mathbb{F}_i] = XE[Y|\mathbb{F}_i]$$

Since $X_0,X_1,X_2,\dots,X_n$ is martingale and let $Y_i = X_i - X_{i-1}$ for $i = 1,2,\dots,t$.

$$\begin{aligned} E[Y_i|X_0,X_1,X_2,\dots,X_{i-1}] &= E[X_i - X_{i-1}|X_0,X_1,\dots,X_{i-1}] \\ &= E[X_i|X_0,X_1,\dots,X_{i-1}] - X_i = 0 \end{aligned}$$

Now consider, $Y_i = -c_i\frac{1-\frac{Y_i}{c_i}}{2} + c_i\frac{1+\frac{Y_i}{c_i}}{2}$

Using the convexity of the function $e^{\alpha Y_i}$

$$\begin{aligned} e^{\alpha Y_i} &\le \frac{1-\frac{Y_i}{c_i}}{2}e^{-\alpha c_i} + \frac{1+\frac{Y_i}{c_i}}{2}e^{\alpha c_i} \\ &= \frac{e^{\alpha c_i}+e^{-\alpha c_i}}{2} + \frac{Y_i}{2c_i}(e^{\alpha c_i}+e^{-\alpha c_i}) \end{aligned}$$

$$\begin{aligned} E[e^{\alpha Y_i}|X_0,X_1,\dots,X_{i-1}] &\le E\left[\frac{e^{\alpha c_i}+e^{-\alpha c_i}}{2} + \frac{Y_i}{2c_i}(e^{\alpha c_i}+e^{-\alpha c_i})\middle|X_0,X_1,\dots,X_{i-1}\right] \\ &= \frac{e^{\alpha c_i}+e^{-\alpha c_i}}{2} \end{aligned}$$

$\le e^{(\alpha c_i)^2/2}$ Using Taylor series

$$\Rightarrow E[e^{\alpha Y_i}|X_0,X_1,\dots,X_{i-1}] \le e^{(\alpha c_i)^2/2}$$

$$\begin{aligned} E\left[e^{\alpha(X_t-X_0)}\right] &= E[\prod_{i=1}^{t} e^{\alpha Y_i}] \\ &= E[\prod_{i=1}^{t-1} e^{\alpha Y_i}]\,E[e^{\alpha Y_i}|X_0,X_1,\dots,X_{i-1}] \\ &\le E[\prod_{i=1}^{t-1} e^{\alpha Y_i}]\,e^{(\alpha c_i)^2/2} \end{aligned}$$

$$\leq e^{\alpha^2 \sum_{i=1}^{t} {c_i}^2/2}$$

$\leq e^{\alpha^2 tc^2/2}$ As $c_i = c \; \forall \; i$

$$\Pr[X_t - X_0 \geq l] = \Pr[e^{\alpha(X_t - X_0)} \geq e^{\alpha l}]$$
$$\leq \frac{E[e^{\alpha(X_t - X_0)}]}{e^{\alpha l}}$$
$$\leq e^{\alpha^2 tc^2/2 - \alpha t}$$
$$= e^{-l^2/(2tc^2)} \; where \; \alpha = \frac{l}{tc^2} \Rightarrow \Pr[|X_t - X_0| \geq l] \leq 2\exp(-\frac{l^2}{2tc^2})$$

III. A SURVEY OF RANDOM GRAPH MODEL FOR WEB MODELING

Random graphs

Random graph is a graph which is generated by some random process. Two basic models for generating random graphs are $G_{n,N}$ and $G_{n,p}$ models. As we know there are many NP- hard computational problems defined on graphs: Hamiltonian cycle, independent set, Vertex cover, and so forth. One question worth asking is whether these problems are hard for most inputs or just for a relatively small fraction of all graphs. Random graph models provide a probabilistic setting for studying such questions.

$G_{n,N}$ Model of Random Graph

In $G_{n,N}$ model of random graph n is the count of vertices and N is number of edges in the graph. We consider all undirected graphs on n vertices with exactly N edges. Since, there are n vertices so possible number of edges are $\binom{n}{2}$. Out of $\binom{n}{2}$ many possible edges graph has N edges, this can be selected in $\binom{\binom{n}{2}}{N}$ manyways. So total possible graphs are $\binom{\binom{n}{2}}{N}$.

Generating graph using $G_{n,N}$ model

One way to generate a graph from graphs in $G_{n,N}$ is to start with no edges. Choose one of the $\binom{n}{2}$ possible edges uniformly at random and add it to the edges in the graph. Now choose one of the remaining $\binom{n}{2} - 1$ possible edges independently and uniformly at random and add to the graph. Similarly, continue choosing one of the remaining unchosen edges independently at random until there are N edges.

$G_{n,p}$ Model of Random Graph

In $G_{n,p}$ model of random graph n is the count of vertices and p is the probability of selection of an edge. We consider all undirected graphs on n distinct vertices $v_1, v_2, \ldots, v_n$. A graph with a given set of m edges has probability $p^m(1 - p)^{\binom{n}{2} - m}$.

<u>Generating graph using $G_{n,p}$ model</u>

One way to generate a random graph in $G_{n,p}$ is to consider each of the $\binom{n}{2}$ possible edges in some order and then independently add each edge to the graph with probability p. i.e. Corresponding to each edge out of $\binom{n}{2}$ throw a coin biased with outcome head with probability p. If outcome is head add the edge to the graph, don't add otherwise. The expected number of edges in the graph is therefore $\binom{n}{2}$p, and each vertex has expected degree (n-1) p.

The $G_{n,N}$ and $G_{n,p}$ models are related: when $p = N/\binom{n}{2}$. The $G_{n,p}$ and $G_{n,N}$ are static models. In these models, graph G = (V,E) is not changing with time. Sometimes we need a graph model which is time evolving $G_t = (V_t, E_t)$.

Evolving Model

Evolving Model of graph covers evolving nature of it. There are two characteristic functions of evolving graph model [5] first gives the number of vertex added at time t+1, whereas second gives set of edges added at time t+1, these are called characteristic function of evolving graph model.

Characteristic functions of evolving graph model $f_v(V_t,t)$ and $f_e(f_t,G_t,t)$. $f_v(V_t,t)$ gives number of vertex added at time t + 1 and $f_e(f_t,G_t,t)$ gives the set of edges added at time t + 1.

$$V_{t+1} = V_t + f_v(V_t,t)$$
$$E_{t+1} = E_t \cup f_e(f_t,G_t,t)$$

The evolving graph model is completely characterized by f_v and f_e.

Linear growth copying model

The Linear growth copying model is evolving model. In each time interval one vertex is added. This is linear growth, because at timestep t, a single vertex arrives and may link to any of the first t-1 vertices. This model describes web graph easily. Web graph is time evolving in nature where at timestep one web page may arrive and is connected to previous one, or a page may be deleted.

Generating graph using linear growth copying model [5]

In each time interval one vertex is added. Edges are added in following way: for a new vertex *u* out-degree *d* is constant. The i^{th} out-degree is decided as- With probability α destination is chosen uniformly from V_t and with remaining probability destination is chosen as the i^{th} outlink of prototype vertex p. Where α is copy factor α∈(0, 1) and the out-degree is d ≥ 1 and $f_v(V_t, t) = 1$. The prototype vertex is chosen once in advance.If $N_{t,i}$ represents number of vertices having in-degree i at time t, for simplicity *d* = 1 and *i* = 0.

In-degree $N_{t,0}$ distribution

Theorem 6. $E[Nt,0 \mid Nt\text{-}k,0] = Nt\text{-}k_{,0}S_{k,0} + \sum_{j=0}^{k-1} Sj,0$.

Proof

$E[Nt,0 \mid Nt\text{-}k,0] = 1 + (1-\frac{\alpha}{t-1})+(1-\frac{\alpha}{t-1})(1-\frac{\alpha}{t-2})+\ldots$
$+(1-\frac{\alpha}{t-1})(1-\frac{\alpha}{t-2})\ldots.(1-\frac{\alpha}{t-k})N_{t-k,0}$

$\mathbf{S_{0,0}=1}$

$S_{k,0}=S_{k-1,0}(1-\frac{\alpha}{t-k})$

$E[N_{t,0} \mid N_{t-k,0}] =N_{t-k,0}S_{k,0}+\sum_{j=0}^{k-1} sj,0$

$S_{k,0}= S_{k-1,0}(1-\frac{\alpha}{t-k})$

$E[N_{t,0} \mid N_{t-k,0}]=N_{t-k,0}S_{k,0}+\sum_{j=0}^{k-1} Sj,0$

t-k=1

$E[N_{t,0} \mid N_{1,0}]=N_{1,0}S_{t-1,0}+\sum_{j=0}^{t-2} Sj,0$

$E[N_{t,0}]=\sum_{j=0}^{t-1} Sj,0$ (As $N_{1,0}$=1)

Theorem 7. $E[N_{t,0} \mid N_{t-k,0}] - E[N_{t,0} \mid N_{t-k+1),0}] \mid \leq 1$

Proof.

$\mid E[N_{t,0} \mid N_{t-k,0}] - E[N_{t,0} \mid N_{t-(k+1),0}] \mid = \mid N_{t-k,0}S_{k,0} +\sum_{j=0}^{k-1} Sj\ ,0 - N_{t-(k-1),0}S_{k+1,0}-\sum_{j=0}^{k} Sj\ ,0 \mid$

Case 1:

$(N_{t-k,0}=N_{t-(k+1),0}+1)$

$= \mid N_{t-(k+1),0}S_{k,0} + S_{k,0} - N_{t-(k+1),0}S_{k,0} (1 -\frac{\alpha}{t-(k+1)})-S_{k,0} \mid$

$=\mid S_{k,0}-S_{k,0}(1-\frac{\alpha Nt\ -(k+1),0}{t-(k+1)}) \mid \leq 1$

$(S_{k,o}(1-\frac{\alpha Nt\ -(k+1,0)}{t-(k+1)})\leq 1)$

From equations it is probability which is always ≤ 1 and $S_{k,0} \leq 1$.

Case 2.

$(N_{t-k}=N_{t-(k+1),0})$

$=\mid N_{t-(k+1),0}S_{k,0}-N_{t-(k+1),0}S_{k,0}(1-\frac{\alpha}{t-(k+1)})-S_{k,0} \mid$

$=\mid S_{k,o}(1-\frac{\alpha Nt\ -(k+1),0}{t-(k+1)}) \mid \leq 1$

$(S_{k,0}(1-\frac{\alpha Nt\ -(k+1),0}{t-(k+1)})\leq 1)$

From equations ,it is probability which is always ≤ 1 and as $S_{k,0} \leq 1$

$\mid E[N_{t,0} \mid N_{t-k,0}] - E[N_{t,0} \mid N_{t-(k+1),0}] \mid \leq 1$

This result is an improvement compared to the result given in the paper [5]

$\mid E[N_{t,0} \mid N_{t-k,0}] - E[N_{t,0} \mid N_{t-(k+1),0}] \mid \leq 2$

Theorem 8. $E[N_{t,0}] = \frac{t}{1+\alpha}$.

Proof. From equation

$$S_{k,0} = S_{k-1,0}\left(1-\frac{\alpha}{t-k}\right)$$

$$S_{k,0} - S_{k-1,0} = -S_{k-1,0}\frac{\alpha}{t-k}$$

$$(S_{k,0}-S_{k-1,0})(t-k) = -\propto S_{k-1,0}$$

$(t-k)(S_{k-1,0} - S_{k,0}) = \propto S_{k-1,0}$

$(t-(k-1))(S_{k-1,0} - S_{k,0}) = (1+\alpha) S_{k-1,0}$

$(1+\alpha) S_{k-1,0} = (t-k) S_{k,0} - (t-(k+1)) S_{k+1,0}$

$(1+\alpha) S_{0,0} = tS_{0,0} - (t-1) S_{1,0}$

$(1+\alpha) S_{1,0} = (t-1) S_{1,0} - (t-2) S_{2,0}$

$(1+\alpha) S_{2,0} = (t-2) S_{2,0} - (t-3) S_{3,0}$

:

$(1+\alpha) S_{t-2,0} = 2S_{t-2,0} - S_{t-1,0}$

$\underline{(1+\alpha) S_{t-1,0} = S_{t-1,0} - (t-t) S_{t,0}}$

$(1+\alpha)\sum_{j=0}^{t-1} S_{j,0} = S_{0,0}$

$$\Rightarrow E[N_{t,0}] = \frac{t}{1+\alpha}$$

This result is an improvement compared to the result given in the paper [5]

$\frac{t+\alpha}{1+\alpha} - \alpha^2 \ln t \le E[N_{t,0}] \le \frac{t+\alpha}{1+\alpha}$

Theorem 9. $\Pr[|N_{t,0} - E[N_{t,0}]| \ge l] \le 2\exp\left(-\frac{l^2}{2t}\right)$

Proof. Azuma's Inequality

Let $X_0, X_1, X_2, \ldots$ be a martingale sequence such that for all k

$$|X_k - X_{k-1}| \le c$$

where c may depend on k.

Then for all $t \ge 0$ and any $\lambda > 0$

$$\Pr[|X_t - X_0| \ge l] \le \exp\left(-\frac{l^2}{2t\,c^2}\right)$$

From theorem 7

$$|E[N_{t,0}|N_{t-k,0}] - E[N_{t,0}|N_{t-(k+1),0}]| \le 1$$

$$\Rightarrow \Pr[|N_{t,0} - E[N_{t,0}]| \ge l] \le \exp\left(-\frac{l^2}{2t}\right)$$

This result is an improvement compared to the result given in the paper[5]

$$\Pr[|N_{t,0} - E[N_{t,0}]| \ge l] \le \exp\left(-\frac{l^2}{4t}\right)$$

$$E[N_{t,0}] = \sum_{j=0}^{t-1} S_{j,0} = \frac{t}{1+\alpha}$$

$P_t = \lim_{t\to\infty} \frac{E[N_{t,0}]}{t} = \frac{1}{1+\alpha}$.

IV. DISTRIBUTIONS

In-degree $N_{t,i}$ distribution

$N_{t,i}$ represents number of vertices having in- degree i at time tfor simplicity d = 1 *Hence* $N_{t,i}$ *is*

$$\begin{cases} N_{t-1,i}-1 & wp \quad \frac{\alpha N_{t-1,i}}{t-1}+\frac{(1-\alpha)N_{t-1,i}}{t-1} \\ N_{t-1,i}+1 & wp \quad \frac{\alpha N_{t-1,i-1}}{t-1}+\frac{(1-\alpha)(i-1)N_{t-1,i-1}}{t-1} \\ N_{t-1,i} & otherwise \end{cases}$$

define$X_{t-k}=E[N_{t,i}|N_{t-k,i}N_{*,i-1}]$where $N_{*,i-1}=N_{0,i-1},N_{1,i-1},N_{2,i-1},N_{3,i-1},\dots,N_{t-1,i-1}$for $0\le t-k\le t$

The sequence $X_0,X_1,X_2,\dots X_t$ forms Doob's martingale [3]

$$E\left[N_{t,i}|N_{t-k,i}N_{*,i-1}\right]=N_{t-k,i}S_{k,i}+\sum_{j=0}^{k-1}S_{j,i}F_{j+1,i-1}$$

where

$$S_{0,i}=1$$
$$S_{k,i}=S_{k-1,i}\left(1-\frac{\alpha}{t-k}-\frac{(1-\alpha)i}{t-k}\right)$$
$$F_{k,i-1}=\frac{\alpha+(1-\alpha)(i-1)}{t-k}$$

Proof.

$$E[N_{t,0}|N_{t-1,i}=\beta,N_{t-1,i-1}=\gamma]$$
$$=(\beta-1)\frac{\alpha\beta}{t-1}+\frac{(1-\alpha)i\beta}{t-1}+(\beta+1)\frac{\alpha\gamma+(1-\alpha)(i-1)\gamma}{t-1}$$
$$+\beta-\left(\frac{\alpha\beta}{t-1}+\frac{(1-\alpha)i\beta}{t-1}\right)-\frac{\alpha\gamma+(1-\alpha)(i-1)\gamma}{t-1}$$
$$=\left(1-\frac{\alpha}{t-1}-\frac{(1-\alpha)i}{t-1}\right)\beta+\frac{\alpha+(1-\alpha)(i-1)}{t-1}\gamma$$
$$\Rightarrow E\left[N_{t,i}|N_{t-1,i}N_{t-1,i-1}\right]=\left(1-\frac{\alpha}{t-1}-\frac{(1-\alpha)i}{t-1}\right)|N_{t-1,i}+\frac{\alpha+(1-\alpha)(i-1)}{t-1}N_{t-1,i-1}$$
$$E\left[N_{t,i}|N_{t-2,i},N_{t-2,i-1},N_{t-1,i-1}\right]=E\left[E\left[N_{t,i}|N_{t-1,i}N_{t-1,i-1}\right]|N_{t-2,i},N_{t-2,i-1},N_{t-1,i-1}\right]$$
$$=\left(1-\frac{\alpha}{t-1}-\frac{(1-\alpha)i}{t-1}\right)E\left[N_{t-1,i}|N_{t-2,i},N_{t-2,i-1},N_{t-1,i-1}\right]$$
$$+\frac{\alpha+(1-\alpha)(i-1)}{t-1}E\left[N_{t-1,i-1}|N_{t-2,i},N_{t-2,i-1},N_{t-1,i-1}\right]$$
$$E\left[N_{t,i}|N_{t-2,i},N_{t-2,i-1}\right]=\left(1-\frac{\alpha}{t-1}-\frac{(1-\alpha)i}{t-1}\right)\{$$
$$\left(1-\frac{\alpha}{t-2}-\frac{(1-\alpha)i}{t-2}\right)N_{t-2,i}$$
$$+\frac{\alpha+(1-\alpha)(i-1)}{t-2}N_{t-2,i-1}\}+\frac{\alpha+(1-\alpha)(i-1)}{t-1}N_{t-1,i-1}$$
$$\because E\left[N_{t-1,i-1}|N_{t-2,i},N_{t-2,i-1},N_{t-1,i-1}\right]=N_{t-1,i-1}\ and$$
$$E\left[N_{t-1,i-1}|N_{t-2,i},N_{t-2,i-1},N_{t-1,i-1}\right]=E\left[N_{t-1,i}|N_{t-2,i},N_{t-2,i-1}\right]$$
$$E\left[N_{t,i}|N_{t-3,i},N_{t-3,i-1}N_{t-2,i},N_{t-2,i-1},N_{t-1,i-1}\right]$$
$$=E\left[\begin{matrix}E\left[N_{t,i}|N_{t-2,i},N_{t-2,i-1},N_{t-1,i-1}\right]|N_{t-3,i},N_{t-3,i-1}N_{t-2,i},\\ N_{t-2,i-1},N_{t-1,i-1}\end{matrix}\right]=\left(1-\frac{\alpha}{t-1}-\frac{(1-\alpha)i}{t-1}\right)$$
$$\{\left(1-\frac{\alpha}{t-2}-\frac{(1-\alpha)i}{t-2}\right)E\left[N_{t-2,i}|N_{t-3,i},N_{t-3,i-1}N_{t-2,i},N_{t-2,i-1},N_{t-1,i-1}\right]$$
$$+\frac{\alpha+(1-\alpha)(i-1)}{t-2}E\left[\begin{matrix}N_{t-2,i-1}|N_{t-3,i},N_{t-3,i-1}N_{t-2,i}\\ ,N_{t-2,i-1},N_{t-1,i-1}\end{matrix}\right]\}$$
$$+\frac{\alpha+(1-\alpha)(i-1)}{t-1}E\left[\begin{matrix}N_{t-1,i-1}|N_{t-3,i},N_{t-3,i-1}N_{t-2,i},\\ N_{t-2,i-1},N_{t-1,i-1}\end{matrix}\right]$$

$$\Rightarrow E\left[N_{t,i} | N_{t-k,i} N_{*,i-1}\right] = N_{t-k,i} S_{k,i} + \sum_{j=0}^{k-1} S_{j,i} F_{j+1,i-1}$$

where

$$\boldsymbol{S_{0,i} = 1}$$

$$S_{k,i} = S_{k-1,i} (1 - \frac{\alpha}{t-k} - \frac{(1-\alpha)i}{t-k})$$

$$F_{k,i-1} = \frac{\alpha + (1-\alpha)(i-1)}{t-k} \quad for \quad k \geq 1$$

V. SURVEY INFERENCE

The synthesis

$$\boldsymbol{|E\left[N_{t,i} | N_{t-k,i} N_{t-k,i-1}\right] = E\left[N_{t,i} | N_{t-(k+1),i} N_{t-(k+1),i-1}\right] | \leq 2}$$

Proof.

$$E\left[N_{t,i} | N_{t-k,i} N_{t-k,i-1}\right] = E\left[N_{t,i} | N_{t-(k+1),i} N_{t-(k+1),i-1}\right] |$$

$$= N_{t-k,i} S_{k,i} + \sum_{j=0}^{k-1} S_{j,i} F_{j+1,i-1} - N_{t-(k+1),i} S_{k+1,i} + \sum_{j=0}^{k} S_{j,i} F_{j+1,i-1}$$

$$= N_{t-k,i} S_{k,i} - N_{t-(k+1),i} S_{k+1,i} - S_{k,i} F_{k+1,i-1}$$

Case 1:

$$(N_{t-k,i} = N_{t-(k+1),i} - 1)$$

$$= |N_{t-(k+1),i} S_{k,i} - S_{k,i} - N_{t-(k+1),i} S_{k,i} (1 - \frac{\alpha}{t-(k+1)} - \frac{(1-\alpha)i}{t-(k+1)})$$

$$-S_{k,i} \left(\frac{\alpha + (1-\alpha)(i-1)}{t-(k+1)}\right) N_{t-(k+1),i-1}|$$

$$= | - S_{k,i} + S_{k,i}\{(\frac{\alpha N_{t-(k+1),i}}{t-(k+1)} + \frac{(1-\alpha)N_{t-(k+1),i}}{t-(k+1)})$$

$$- \left(\frac{\alpha + (1-\alpha)(i-1)}{t-(k+1)}\right) N_{t-(k+1),i-1}\}|$$

$$\leq 2$$

$$(\because \left(\frac{\alpha + (1-\alpha)(i-1)}{t-(k+1)}\right) N_{t-(k+1),i-1} \leq 1)$$

$$(\because \left(\frac{\alpha N_{t-(k+1),i}}{t-(k+1)} + \frac{(1-\alpha)N_{t-(k+1),i}}{t-(k+1)}\right) \leq 1)$$

Bound for Case 1 is same as that of paper [3].

Case 2:

$$(N_{t-k,i} = N_{t-(k+1),i} + 1)$$

$$= |N_{t-(k+1),i} S_{k,i} + S_{k,i}$$

$$-N_{t-(k+1),i} S_{k,i} (1 - \frac{\alpha}{t-(k+1)} - \frac{(1-\alpha)i}{t-(k+1)})$$

$$-S_{k,i} \left(\frac{\alpha + (1-\alpha)(i-1)}{t-(k+1)}\right) N_{t-(k+1),i-1}|$$

$$= |S_{k,i} + S_{k,i}\{(\frac{\alpha N_{t-(k+1),i}}{t-(k+1)} + \frac{(1-\alpha)N_{t-(k+1),i}}{t-(k+1)})$$

$$-(1 - \left(\frac{\alpha+(1-\alpha)(i-1)}{t-(k+1)}\right) N_{t-(k+1),i-1})\} \leq 2$$

Bound for Case 2 is same as that of paper [3]

Case 3:

$$(\because N_{t-k,i} = N_{t-(k+1),i})$$

$$= N_{t-(k+1),i} S_{k,i} - N_{t-(k+1),i} S_{k,i} (1 - \frac{\alpha}{t-(k+1)} - \frac{(1-\alpha)i}{t-(k+1)})$$

$$-S_{k,i}\left(\frac{\alpha+(1-\alpha)(i-1)}{t-(k+1)}\right)N_{t-(k+1),i-1}$$
$$=S_{k,i}\left(\frac{\alpha N_{t-(k+1),i}}{t-(k+1)}+\frac{(1-\alpha)N_{t-(k+1),i}}{t-(k+1)}\right)$$
$$-S_{k,i}\left(\frac{\alpha+(1-\alpha)(i-1)}{t-(k+1)}\right)N_{t-(k+1),i-1}$$
$$\leq 1$$

Bound for Case 3 obtained here is an improvement compared to the result givenin the paper [5]

Azuma's Inequality:

Let $X_0, X_1, X_2, \ldots$be a martingale sequence such that for all k

$$|X_k - X_{k-1}| \leq c$$

where$c$may depend on$k$.
Then for all t≥0 and any λ>0

$$\because |E[N_{t,i}|N_{t-k,i}N_{t-k,i-1}] = E[N_{t,i}|N_{t-(k+1),i}N_{t-(k+1),i-1}]| \leq 2$$
$$\Rightarrow \Pr[|N_{t,i} - E[N_{t,i}| \geq l] \leq \exp(-\frac{l^2}{4t})$$

VI. CONCLUSION

This report studies random graph models using probabilistic techniques.Concentration is made over the paper Stochastic Model for Web graph by RaviKumar et.al. [5]. during the proof of the results given in the paper some moreprecise results have been found. Example- for i = 0 report proves $|E[N_{t,0}|N_{t-k,0}]$- $E[N_{t,0}|N_{t-(k+1),0}]| \leq 1$whereas in the paper it is $|E[N_{t,0}|N_{t-k,0}]$- $E[N_{t,0}|N_{t-(k+1),0}]| \leq 2$. It is also proved that the $E[N_{t,0}] = \frac{t}{1+\alpha}$. For general value of i, report proves that$|E[N_{t,i}|N_{t-k,i}N_{t-k,i-1}]$ - $|E[N_{t,i}|N_{t-(k-1),i}N_{t-(k-1),i-1}]| \leq 1$for $N_{t-k,I} = N_{t-(k+1),i}$whereas result of the paper is $|E[N_{t,i}|N_{t-k,i}N_{t-k,i-1}]$ - $|E[N_{t,i}|N_{t-(k-1),i}N_{t-(k-1),i-1}]| \leq 2$. In future one can find improvement in $|E[N_{t,i}|N_{t-k,i}N_{t-k,i-1}]$ - $|E[N_{t,i}|N_{t-(k-1),i}N_{t-(k-1),i-1}]|$ for general value of $N_{t-k,i}$.

VII. APPENDIX

REFERENCES

[1] W. Aiello, F. Chung, and L. Lu. A random graph model for massive graphs. 2000.
[2] B. Bollobas. Random graphs. 1985.
[3] J.L.Doob. What is a martingale. The American Mathematical Monthly Vol 78, No.5 (May 1971), pp. 451-46, 1971.
[4] J. Kleinberg, S. R. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins. The web as a graph: Measurements, models and methods. 1999.
[5] Ravi Kumar, PrabhakarRaghavan, Sridhar Rajagopalan, D. Sivakumar, Andrew Tomkins, and Eli Upfal. Stochastic models for the web graph. 2000.
[6] Michael Mitzenmacher and Eli Upfal. Probability and computing.
[7] Rajeev Motawani and PrabhakarRaghavan. Randomized algorithm. 2009.
[8] Sheldom M. Ross. Probability models for computer science. 2006.