Scaling laws for file dissemination in P2P networks with random contacts

Rudesindo Núñez-Queija*† and Balakrishna Prabhu* ‡§
*CWI, Kruislaan 413, 1098 SJ Amsterdam, The Netherlands
† TNO Information and Communication Technology, Delft, The Netherlands
‡ LAAS-CNRS, 7 Av. du Colonel Roche, 31077 Toulouse Cedex 4, France
Email: sindo@cwi.nl and bjprabhu@laas.fr

***Abstract*—In this paper we obtain the scaling law for the mean broadcast time of a file in a P2P network with an initial population of N nodes. In the model, at Poisson rate λ a node initiates a contact with another node chosen uniformly at random. This contact is said to be successful if the contacted node possesses the file, in which case the initiator downloads the file and can later upload it to other nodes. In a network with altruistic nodes (i.e., nodes do not leave the network) we show that the mean broadcast time is $O(\log(N))$. In a network with free-riding nodes, our main result shows that a $O(\log(N))$ mean broadcast time can be achieved if nodes remain connected to the network for the duration of at least one more contact after downloading the file, otherwise a significantly worse $O(N)$ time is required to broadcast the file.**

I. Introduction

Traffic measurements in the Internet suggest that Peer-to-Peer (P2P) networks are becoming increasingly popular among Internet users for sharing and distributing files. The salient features of a P2P architecture are the vast possible improvements in scalability and robustness compared to the traditional client-server architecture. In the best-case scenario, a P2P network can broadcast a file in a time which scales only logarithmically with the number of nodes in the network, which compares favourably with the linear scaling for a client-server network. This vast improvement in the distribution time can be explained as follows. After downloading the file, a client node acts as a server and uploads the file to other client nodes. Thus, the service capacity of the network actually increases with the number of the nodes in the network. The presence of several simultaneous servers in the network significantly reduces the vulnerability of the file distribution process to attacks on the central server.

Although the P2P architecture is very promising in terms of scalability, there are several factors which are critical to achieving the promised performance gains. The foremost factor is the willingness of each client node to become a server node. A failure on the part of client nodes to do so (also called *free-riding*) would impact both scalability and robustness. As a simple example, if each client node departs immediately after having downloaded the entire file, the network will behave as a client-server network with the broadcast time scaling linearly in the number of client nodes. Thus, the impact of free-riding (i.e., downloading but not uploading) on the broadcast time needs a detailed investigation. Another factor which is critical to achieving the performance gains is the connectivity of the underlying network graph. Again as a simple example, if the network would be configured in a star topology then again the time to broadcast would be linear in the number of nodes which is significantly worse than the logarithmic scaling possible on a hypercube topology [4].

This research is financially supported by the SCALP project of the Network of Excellence EURO-FGI and by the Dutch Bsik/BRICKS project. The second author wishes to acknowledge the support of CWI, the Department of Mathematics and Computer Science at Eindhoven University of Technology, and EURANDOM, where this work was carried out.

A detailed study of the impact of these factors on the performance of a P2P network is thus essential in obtaining conditions under which P2P architecture can outperform the client-server architecture. In this paper, we study a closed P2P network in which N client nodes and one seed node form a fully connected file sharing network. The purpose of this network is to broadcast the file which is available at the seed node. Each node (except the seed node) can leave the network after downloading the file. The model described above is suited to study the behaviour of P2P networks when subjected to *flash-crowds*, i.e., a large population of nodes joins the network in a very short interval of time [2]. One of the main performance measures in such networks is the time required to broadcast the file. The focus of this paper is to study the impact of free-riding on the mean broadcast time. Our main result states that a $O(\log(N))$ mean broadcast time is achievable in P2P networks with free-riding provided that nodes stay long enough in the network after having downloaded the file, otherwise a significantly worse $O(N)$ time is required, thereby implying a phase transition phenomenon for the scaling law of the mean broadcast time.

A. Related work

The availability of free P2P software such as BitTorrent [1] has contributed significantly to the increased popularity of P2P networks among Internet users, and has also motivated research in several aspects of the P2P networks. The BitTorrent P2P algorithm achieves a significant improvement in performance by dividing the file into several chunks. Instead of downloading a large file from one server, nodes can download smaller chunks from different servers. A file download is said to be complete when a node has downloaded all the corresponding chunks. Previously, low bandwidth nodes were reluctant to upload because of large file sizes, and thus reluctant to participate in P2P networks. However, breaking the file into smaller chunks provides such users an incentive to upload data and join a P2P network. In [4], the authors studied the problem of the optimal broadcast of a set of C messages to N nodes over a complete graph in a *deterministic* setting. They showed that the optimal broadcast time is $O(C+\log_2(N))$. In our present paper, we give an analogous result in a stochastic setting for the one chunk case and with the more realistic assumption of nodes being able to leave the network.

In general, the analysis of P2P network in a stochastic setting (i.e., random node arrival and node departures) is too complex to permit an exact analysis. Hence approximate models have been constructed to obtain some insights into the performance of P2P networks. For example, using a fluid model Qiu and Srikant [6] have studied the behaviour of the number of servers and clients in a BitTorrent network with external arrivals and node departures. The emphasis is on studying the number of servers and clients in the equilibrium state. In [3], the authors generalized the above model to be able to study the spread of chunks within networks. One of their results shows that chunk selection policies (like *rarest first* or *random selection*) have negligible impact on the performance of a P2P network. In practice, arrivals to a network may not occur at a constant rate, and the so-called *flash crowd* phenomenon has often been observed [2]. For example, the latest version of a popular software is solicited by a large number of users (a flash crowd) close to the release date. Usually, the interest in this version may taper off as time progresses, and the critical period of operation is during the first few days when the interest is large. We note that the interest may increase again when a new version of the software is released, for example. Unlike the above mentioned work, our objective in this paper is to characterize the mean *broadcast time* in a *closed* network. In that respect, our work is an extension of [4] to stochastic setting with *free-riding*. However, the analytical tools (Markov chains and fluid limits) are similar to those in [6] and [3].

As a first step, we present the analysis for the one chunk case, i.e., the file is not divided as in BitTorrent. From the insights obtained using this model, we intend to extend this analysis to the multiple chunk case and for different network topologies.

The rest of the paper is organised as follows. In section II, we describe the model, give the assumptions, and formulate the problem in terms of the input parameters. The analysis for a network without free-riding is presented in section III. In section IV our main result on the mean broadcast time in a network with free-riding users is derived. Using simulations, similar results for general values of C are given in section V. Finally, we conclude with possible research directions in section VI.

II. Problem Formulation

Consider a population of N nodes who want to download a file which is available at the seed node at time 0. We assume that the underlying network topology is fully connected, and that a node, which is present in the network and has the file, is willing to upload the file to other nodes. In order to download a file, a node initiates a contact with another node chosen uniformly at random among the existing nodes. These contacts are initiated at Poisson rate λ. If the contacted node has the file then the file transfer is assumed to take place in a time which is negligible compared to the mean time between contacts. This model of a contact process for file dissemination is based on the one analysed in [3] and [5].

In order to model the impatient behaviour of nodes in a real network, we shall assume that, after having downloaded the file, a node leaves the network at a Poisson rate ν. The case $\nu = 0$ corresponds to altruistic nodes who remain in the network for the duration of the broadcast whereas the case $\nu = \infty$ corresponds to nodes who leave the network immediately after downloading the file. Finally, we shall assume that the seed node remains in the network for the duration of the broadcast. This assumption guarantees that all the nodes will be able to download the file eventually. One could possibly study the number of unsuccessful nodes if the seed node also had the possibility of leaving the network. Such an analysis could give clues to the vulnerability of the network to malicious attacks on the seed node.

Given the above setting, our main interest in this paper is to study the impact of the departure rate, ν, on the mean time to broadcast the file to all N nodes. Intuitively, a higher departure rate of the nodes would translate into fewer servers present in the network which would then increase the mean broadcast time. We shall formalize this intuitive result by showing that, depending on the departure rate, different scaling laws are possible for the mean broadcast time.

III. Mean broadcast time with altruistic nodes ($\nu = 0$)

We first take a look at the case $\nu = 0$. Through this analysis we expect to obtain a lower bound on the mean broadcast time for $\nu > 0$. In a deterministic setting when the sequence in which file downloads take place is determined at time 0, file broadcast can be achieved in $O(\log(N))$ time units. We now show that this is also the case in the stochastic contact process model we described earlier. Thus, the mean broadcast time in a random contact based P2P network is of the same order as the optimal broadcast time.

For the case of $\nu = 0$, we shall study the network in discrete time where each time step corresponds to the time between two contacts. Since no nodes leave the network, contacts are initiated at rate $N\lambda$ (we assume that the seed does not initiate any contacts). The mean broadcast time can be obtained by multiplying the mean number of contacts by $(N\lambda)^{-1}$.

Let Y_n denote the number of servers in the network after the nth contact. The dynamics of the process $\{Y_n, n > 0\}$ can be described as follows.

$$Y_{n+1} = \begin{cases} Y_n & \text{w.p.} \quad p(Y_n) \\ Y_n + 1 & \text{w.p.} \quad 1 - p(Y_n) \end{cases},$$

where $p(i) = 1 - \frac{N-i}{N}\frac{i+1}{N}$. The probability $p(i)$ describes the probability of an unsuccessful contact when there are i servers and one seed present in the network.

Let A_i denote the number of contacts made in state i. The random variable A_i is geometrically distributed with

$$\mathrm{P}(A_i = k) = (1 - p(i))p(i)^{k-1},\ k \geq 1.$$

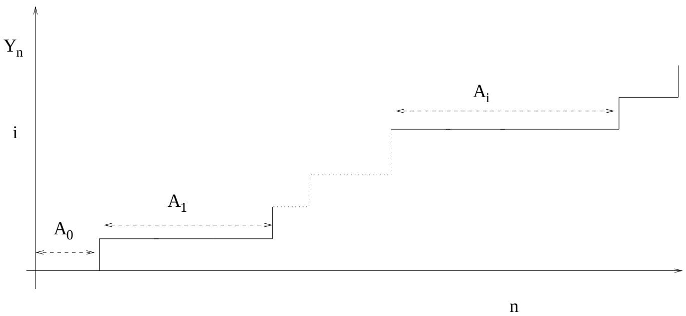

Fig. 1. The relation between S and dynamics of Y

.

Let $S_j = \sum_{i=0}^{j} A_i$. The random variable S_j is the number of contacts needed to distribute the file to $j+1$ nodes. This relation between the processes S_j and Y_n is illustrated in figure 1 from which we can infer that $P(Y_n < j) = P(S_{j-1} > n)$. Since A_j are independent random variables,

$$\begin{aligned} E[S_j] &= \sum_{i=0}^{j} E[A_i], \quad j = 0, 1, ..., N-1, \\ Var[S_j] &= \sum_{i=0}^{j} Var[A_i], \quad j = 0, 1, ..., N-1. \end{aligned}$$

Also, since A_is are geometrically distributed,

$$\begin{aligned} E[A_i] &= \frac{1}{1 - p(i)} \\ &= \frac{N^2}{(i+1)(N-i)}, \\ Var[A_i] &= \frac{p(i)}{(1-p(i))^2} \\ &= \left(1 - \frac{i+1}{N}\left(\frac{N-i}{N}\right)\right)\left(\frac{N^2}{(i+1)(N-i)}\right)^2. \end{aligned}$$

Therefore, the mean number of contacts to broadcast the file (i.e., to distribute the file to N nodes) is

$$\begin{aligned} E[S_{N-1}] &= \sum_{i=0}^{N-1} E[A_i] \\ &= \frac{N^2}{N+1}\left(2\log(N) + o(\log(N))\right). \end{aligned}$$

Let T_j denote the time needed to distribute the file to j nodes. Then,

$$T_j = \sum_{k=0}^{S_{j-1}} \tau_k, \tag{1}$$

where the random variable τ_k denotes the time between the kth and the $(k+1)$th contact. Since $\tau_1, \tau_2, ...$ is a sequence of i.i.d. exponential random variables with mean $(N\lambda)^{-1}$, we can use Wald's lemma and obtain the mean broadcast time as

$$\begin{aligned} E[T_N] &= E[S_{N-1}]E[\tau_1] \\ &= 2\frac{N}{\lambda(N+1)}\log(N) + o(\log(N)). \end{aligned} \tag{2}$$

IV. Mean broadcast time with free riding nodes ($\nu > 0$)

In the previous section we obtained a mean broadcast time of $O(\log(N))$ for $\nu = 0$. For the other extreme case of $\nu = \infty$, we can see that the broadcast time would be $O(N)$ because the seed would be the only server present in the network, and every user will have to download the file from the seed node, which will take $O(N)$ encounters.

In this section we shall obtain the scaling law when $0 < \nu < \infty$, i.e., nodes leave the network at rate ν after downloading the file. Let $Y(t)$ (resp. $X(t)$) denote the number of servers (resp. downloaders) present in the network at time t. The joint process $\{X(t), Y(t)\}_{t \geq 0}$ is a two-dimensional Markov process on $\{0, 1, 2, ..., N\} \times \{0, 1, 2, ..., N\}$ whose dynamics can be described as follows

$$Y(t) \rightarrow \begin{cases} Y(t) + 1 & \text{at rate} \quad \lambda X(t)\frac{Y(t)+1}{X(t)+Y(t)} \\ Y(t) - 1 & \text{at rate} \quad \nu Y(t) \end{cases}, \tag{3}$$

$$X(t) \rightarrow X(t) - 1 \quad \text{at rate} \quad \lambda X(t)\frac{Y(t)+1}{X(t)+Y(t)}, \tag{4}$$

(5)

with $(X(0), Y(0)) = (N, 0)$. The increase in $Y(t)$ only happens when downloaders make a successful contact (the $+1$ in the numerator is due to the presence of the seed). The rate of decrease of $Y(t)$ is $\nu Y(t)$ independent of the number of downloaders.

We now study this process in the large initial population limit i.e., $N \rightarrow \infty$. Let $(x(t), y(t)) \equiv \left(\frac{X(t)}{N}, \frac{Y(t)}{N}\right)$ be the rescaled process. Then, $y(t)$ (resp., $x(t)$) is the fraction of nodes at time t who do (resp., do not) have the file. For $0 < \nu < \infty$, we can write the following fluid equations for the dynamics x and y,

$$\frac{dy}{dt} = -\nu y + \lambda x \frac{y}{x+y}, \tag{6}$$

$$\frac{dx}{dt} = -\lambda x \frac{y}{x+y}. \tag{7}$$

Therefore,

$$\frac{d(x+y)}{dt} = -\nu y. \tag{8}$$

Combining equations 7 and 8, we get

$$\frac{d}{dx}(x+y) = \frac{\nu}{\lambda}\frac{x+y}{x}, \tag{9}$$

which can be solved to obtain

$$x + y = c_0 x^{\frac{\nu}{\lambda}}. \tag{10}$$

We can determine c_0 by noting that $y = 0$ when $x = 1$. Thus, we can characterize the evolution of the fraction of servers as a function of the fraction of downloaders in the network as follows

$$y = -x + x^{\sigma},\ x \in (0, 1), \tag{11}$$

where $\sigma = \nu/\lambda$. In figure IV, we plot solutions of (10) for different values of $\sigma = \nu/\lambda$. As $\nu \to 0$, the solution approaches the line $x+y=1$, which is the case when no nodes leave the network. The solution to the differential equation obtained above is valid only for $\sigma < 1$. For $\sigma > 1$ we obtain negative values for y, which makes the solution infeasible. We now have the following result.

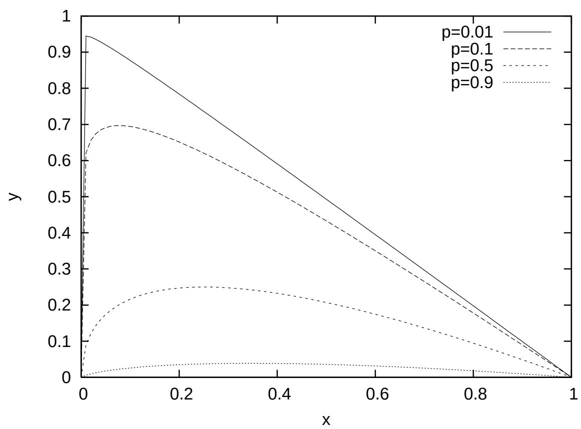

Fig. 2. Solutions of $y = -x + x^{\sigma}$ for various values of σ.

Theorem 4.1: The mean broadcast time for a file in a P2P network with free-riding users scales as

- $O(N)$ if $\sigma > 1$;
- $O(\log(N))$ if $\sigma < 1$.

Thus, there is phase transition in the scaling law at $\sigma = 1$. This suggests that if nodes stay for the duration of one more contact after downloading the file then a significantly improved scaling law for the broadcast time prevails even in the presence of free-riding nodes.

Proof: We first prove that the mean broadcast time is $O(N)$ for $\sigma > 1$. For this case we upper bound $Y(t)$ by another process which is easier to analyse. Let $\{Z(t)\}_{t\geq 0}$ be defined as

$$Z(t) \quad \to \quad \begin{cases} Z(t)+1 & \text{at rate} \quad \lambda Z(t), \\ Z(t)-1 & \text{at rate} \quad \nu Z(t). \end{cases} \qquad (12)$$

For the same initial conditions, $\{Y(t)\}$ is stochastically smaller than $\{Z(t)\}$. For $\sigma > 1$, $\mathrm{P}(Z(t) > N) \to 0$ when $N \to \infty$. Hence, for large N, we can conclude that $Z(t)$ will not reach $O(N)$ and, consequently, Y(t) will remain $o(N)$. From (4), when $X(t)$ is on a linear scale it will decrease at a constant rate. Thus, for large N, the mean time required for $X(t)$ to go from $\alpha_1 N$ to $\alpha_2 N$ will be linear in N. Hence, the mean broadcast time will be $O(N)$.

For $\sigma < 1$, we will follow similar arguments. In order to determine the mean broadcast time, we divide the analysis in three phases. The first phase corresponds to the time required for the number of servers to reach $O(N)$. The second phase begins when both the number of downloaders and the number of servers become $O(N)$. In this phase the dynamics of the rescaled process are governed by the differential equations given in (7) and (8). The second phase ends when the joint process becomes $o(N)$, and we call this the final phase. In this phase the number of dwonloaders eventually goes to zero.

For the time spent in the first phase, we find the mean time required for $Y(t)$ to exceed level $\epsilon_\sigma N$. This level will depend on σ as not all values of $y(t) \in (0,1)$ are feasible for a given σ. First, we find a lower bound for the rate of increase of $Y(t)$. Let γ be the maximal solution of the equation $-x + x^{\sigma} = \epsilon_\sigma$ in $(0,1)$. Then

$$\begin{aligned} \lambda X(t)\frac{Y(t)}{X(t)+Y(t)} &> \lambda Y(t)\frac{X(t)}{X(t)+\epsilon_\sigma} \\ &> \lambda Y(t)\frac{\gamma}{\gamma+\epsilon_\sigma}. \end{aligned}$$

The second inequality follows from the fact that $x/(x+1)$ is an increasing function in x, and that if $Y(t) < \epsilon_\sigma N$ then $X(t) > \gamma N$. We now bound $Y(t)$ by $\hat{Z}(t)$ described by

$$\hat{Z}(t) = \begin{cases} \hat{Z}(t)+1 & \text{at rate} \quad \lambda\frac{\gamma}{\gamma+\epsilon_\sigma}\hat{Z}(t), \\ \hat{Z}(t)-1 & \text{at rate} \quad \nu\hat{Z}(t). \end{cases} \qquad (13)$$

We choose a $\gamma > \sigma^{\frac{1}{1-\sigma}}$ which then determines ϵ_σ. For this choice of γ, $\lambda\frac{\gamma}{\gamma+\epsilon_\sigma} = \lambda\frac{\gamma}{\gamma^{\sigma}} > \lambda\sigma = \nu$. For such a choice of parameters, $\hat{Z}(t)$ and, consequently, $Y(t)$ grow exponentially with time. Hence, the time for $Y(t)$ to reach $\epsilon_\sigma N$, say t_1, is $O(\log(N))$.

For the time spent in the second phase, we first solve (7) to obtain

$$t(x) = \frac{1}{\lambda(1-r)}\log\left(\frac{x+x^{\sigma}}{2x}\right). \qquad (14)$$

From this equation, the time for x to start from a fraction γ and reach a fraction γ^* is a constant independent of N. Hence the time spent in the second phase is $O(1)$.

For the time spent in the third phase, we shall bound the time required for $x(t)$ starting from $x(\tau) = \gamma*$ to reach 0. For a given σ, $y > x$ if $x < \frac{1}{2(1-\sigma)}$. We first fix a $\gamma^* < \frac{1}{2(1-\sigma)}$. For $x < \gamma^*$,

$$\lambda x\frac{y}{x+y} > \frac{1}{2}\lambda x. \qquad (15)$$

Since x is non-increasing, if $x(t_2) < \gamma^*$ then $x(t) < \gamma^*$ and $y(t) > x(t), \forall t > t_2$. Hence, the above inequality will remain valid once x is smaller than γ^*. Let $\{\hat{X}(t)\}$ be described by

$$\hat{X}(t) \quad \to \quad \hat{X}(t)-1 \quad \text{at rate} \quad \tfrac{1}{2}\lambda\hat{X}(t) \; .$$

From this definition, the process $\{X(t)\}$ is stochastically smaller than $\{\hat{X}(t)\}$. Since $\hat{X}(t)$ decreases exponentially, we can conclude that $X(t)$ also decreases to 0 in logarithmic time.

From the above analysis, the time spent in the first phase is upper bounded by $O(\log(N))$, the time spent in the second phase is $O(1)$, and the time spent in the final phase is upper bounded by $O(\log(N))$. Since the time to broadcast cannot be lower than $\log(N)$, we can conclude that the mean broadcast time is $O(\log(N))$ for $\sigma < 1$. ■

V. SIMULATIONS

In this section, we present the results of simulations with C larger than unity. We simulate the same contact model analysed in this paper but with larger number of chunks. In particular, for $C = 10$ and $C = 50$ we shall obtain the mean broadcast time as a function of N and compare the scaling laws for different values of σ.

In figure 3, the mean broadcast time is shown as a function of N for $C = 10$ and two different values of σ smaller than 1. Figure 4 shows the corresponding function for two different values of σ larger than 1.

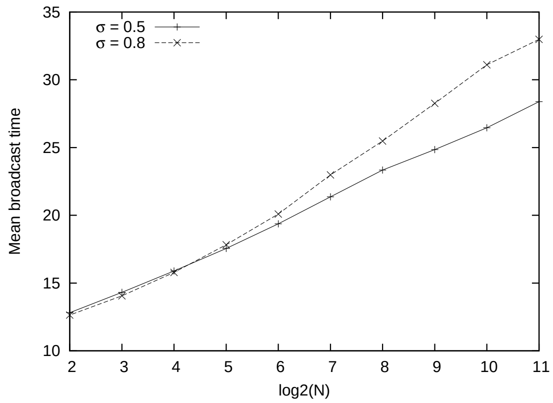

Fig. 3. Mean broadcast time versus $\log_2(N)$. $C = 10$.

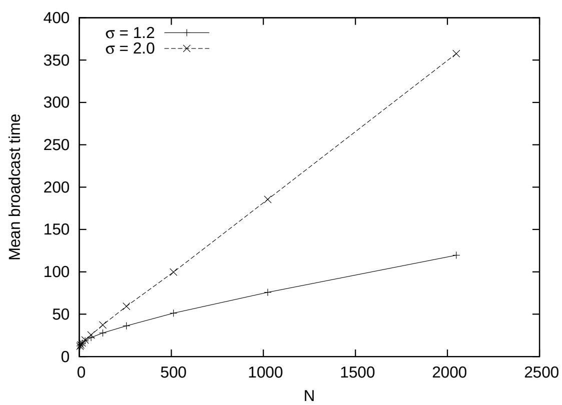

Fig. 4. Mean broadcast time versus N. $C = 10$.

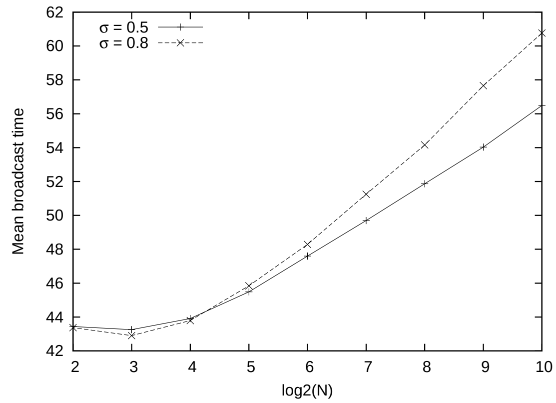

Fig. 5. Mean broadcast time versus $\log_2(N)$. $C = 50$.

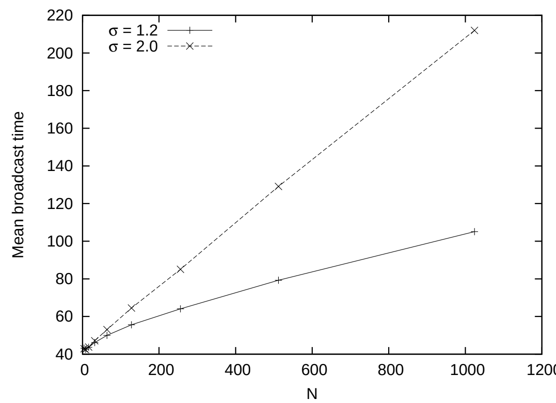

Fig. 6. Mean broadcast time versus N. $C = 50$.

In figures 5 and 6, we plot the mean broadcast time versus N for $C = 50$. From these plots we observe that the phase transition at $\sigma = 1$ equal to appears to be true for larger values of C as well, i.e.. the scaling law is logarithmic scaling when $\sigma < 1$ and linear when $\sigma > 1$. Although, these simulations are not exhaustive, they do point to the plausibility of a strong dependence of the scaling law on the level of user cooperation in multi-chunk dissemination networks as well.

VI. Conclusions and future work

In this paper we quantified the effect of free-riding users on the mean broadcast time of a file in a P2P network. Our main result showed that a logarithmic broadcast time can be achieved if nodes stay in the network for the duration of one more contact, i.e., if they upload the file at least once. Otherwise a significantly worse linear scaling is achieved. Thus, if nodes stay in the network for the duration of one more contact, a random contact based P2P network can broadcast a file in a time which is of the same order as the optimal time.

The main assumptions in this work were on the download times (assumed to be negligible compared to the inter-contact times), graph topology (assumed to be fully-connected) and the number of chunks (assumed to be equal to unity). Our future work will seek to characterize the mean broadcast time in the absence of these assumptions.

References

[1] BitTorrent. [Online]. Available: http://www.bittorrent.com.

[2] Xuan Chen and John Heidemann. Flash crowd mitigation via adaptive admission control based on application-level observations. *ACM Trans. Inter. Tech.*, 5(3):532–569, 2005.

[3] L. Massoulié and M. Vojnović. Coupon Replication Systems. In *Proc. ACM SIGMETRICS*, Banff, Canada, 2005.

[4] J. Mundinger, R. Weber, and G. Weiss. Analysis of peer-to-peer file dissemination. *To appear in Performance Evaluation Review, Special Issue on MAMA 2006*.

[5] I. Norros, B.J. Prabhu, and H. Reittu. Flash crowd in a file sharing system based on random encounters. In *Proc. InterPerf*, Pisa, 2006.

[6] D. Qiu and R. Srikant. Modeling and Performance Analysis of BitTorrent-Like Peer-to-Peer Networks. In *Proc. ACM Sigcomm*, Portland, OR, 2004.