Data Backup Network Formation with Heterogeneous Agents

(Extended Abstract)

Harshit Jain\textsuperscript{a}, Guduru Sai Teja\textsuperscript{a}, Pramod Mane\textsuperscript{a}, Kapil Ahuja\textsuperscript{a} and Nagarajan Krishnamurthy\textsuperscript{b}

Abstract—Social storage systems \cite{1}, \cite{2}, \cite{3} are becoming increasingly popular compared to the existing data backup systems like local, centralized and P2P systems. An endogenously built symmetric social storage model and its aspects like the utility of each agent, bilateral stability, contentment, and efficiency have been extensively discussed in \cite{4}. We include heterogeneity in this model by using the concept of Social Range Matrix from \cite{5}.

Now, each agent is concerned about its perceived utility, which is a linear combination of its utility as well as others utilities (depending upon whether the pair are friends, enemies or do not care about each other). We derive conditions when two agents may want to add or delete a link, and provide an algorithm that checks if a bilaterally stable network is possible or not. Finally, we take some special Social Range Matrices and prove that under certain conditions on network parameters, a bilaterally stable network is unique.

Index Terms—Social Storage, F2F Backup System, Endogenous and Heterogeneous Network Formation, Bilateral Stability.

I. INTRODUCTION AND MOTIVATION

In recent years, social storage (friend-to-friend) is emerging as an alternative to local (e.g., hard disc), online (e.g., cloud), and Peer-to-Peer data backup systems. Social storage allows users to store their data on their social relatives’ (e.g., friends, colleagues, family, etc.) storage devices. Researchers believe that social relationships (which are at the core of social storage) help to mitigate issues like data availability, reliability, and security.

Initial work in this field has primarily focused on developing techniques to exogenously build social storage systems, and performing Quality of Services (QoS) analysis in terms of data reliability and availability in these systems. A most recent study in \cite{4} focuses on explicit data backup partner selection, where agents themselves select their partners. This selection is studied in a strategic setting, and eventually builds a social storage network. They model a utility function, which reveals the benefit an individual receives in a social storage network. Further, they analyze the network by using bilateral stability as a solution concept, where no pair of agents add or delete a link without their mutual consent.

There are several advantages of this approach. First, this approach makes it possible to incorporate user’s strategic behavior (in terms of with whom user wants to form social connections and with whom it does not). Second, this approach helps us to predict the following: which network is likely to emerge; which one is stable (in which no individual has incentives to alter the structure of the network either by forming new or deleting existing social connections); and which one is the best (contended and/ or efficient) from all the participants point of view.

However, the above strategic setting does not consider heterogeneous behavior (e.g., selfless, selfish, etc.) in the network formation. In this paper, we focus on this aspect. This way, we make the model close to a real-world scenario. Although, doing so makes it challenging to deal with the model and as well as predict its outcome.

To achieve the above, we incorporate the concept of Social Range Matrix introduced in \cite{5} while exploring different social relationships between agents. We modify the utility of agents in the social storage model discussed in \cite{4}. Further, we revisit the results regarding bilateral stability of such a model.

II. MODEL

Our model is composed of four components. First is symmetric social storage. Second is the utility (as derived in \cite{4}) of each agent. Third is Social Range Matrix, which captures the social relationship between these agents. Fourth is the perceived utility obtained by combining the second and the third components.

A symmetric social network $g$ is a data backup network consisting of $N$ number of agents and a set of links connecting these agents. A link $(ij) \in g$ represents that the agents $i$ and $j$ are data backup partners, who store their data on each other’s shared storage space. In $g$, pairs of agents share an equal amount of storage space. Agents perform two actions, link addition (represented by $g+(ij)$) and deletion (represented by $g-(ij)$).

The utility function in \cite{4} reveals the cost and benefit that each agent $i$ receives in $g$ and is given by

$$u_i(g) = \beta \cdot (1 - \lambda n_i(g)) - c \cdot n_i(g).$$

This utility function $u_i(g)$ consists of the following parameters: the agent $i$’s neighborhood size $n_i(g)$; the benefit $\beta$ associated with data; the cost $c$ that the agent $i$ incurs to maintain its neighbours; and the probability of disk failure $\lambda$. Note that $c$, $\beta$, and $\lambda$ lie between 0 and 1. The utility function is a combination of two objectives for each agent $i$. The first is to minimize the total cost of the links, which is $c \cdot n_i(g)$. The second is to maximize the expected data backup benefit,
which is $\beta * (1 - \lambda_{i}^{n_{i}(g)})$. Note that from now onwards we will use $n_{i}$ to represent neighbourhood size of agent $i$ in $g$.

In our model, we consider three type of agents. First, where an agent helps to maximize other’s utility (friend). Second, where an agent aims to decrease other’s utility (enemy). Third, where an agent does not care about other’s utility (neutral). Based upon [5], we represent the above kind of social relationships between pairs of agents in a Social Range Matrix $F$. Here, each element $f_{ij}$ denotes the above discussed social relationship between $i$ and $j$. That is,

$$f_{ij}\begin{cases} > 0 & i \text{ and } j \text{ are friend of each other}, \\
< 0 & i \text{ and } j \text{ are enemy of each other}, \\
= 0 & i \text{ and } j \text{ do not care about each other}. 
\end{cases}$$

Also, $j \in$ the set of all agents including $i$ and $f_{ii} > 0$.

Many such interesting matrices are possible, for example, a matrix of all ones mean that all pair of agents are friends, and a matrix of all zeroes except the diagonal elements means weak social links.

In our paper, for the sake of simplicity, we use only three values for $f_{ij}$. That is, 1, -1, and 0 indicating friend, enemy, neutral, respectively. We also consider that agents give more importance to other agents utilities than their own utility, i.e. $f_{ii} < |f_{ij}|$ for all $j$.

Now, agents have different social relationships with others, and hence, their utility not only depends on the structure of the network but also their social relationships with others. We define the new utility (of agent $i$ in $g$) as the perceived utility, which is defined as follows:

$$\tilde{u}_{i}(g) = \sum_{j} f_{ij}u_{j}(g). \quad (1)$$

In this setting, each agent’s objective is to maximize its perceived utility (which takes care of utilities of other agents). Thus, the optimization problem is

$$\max (\tilde{u}_{i}(g)).$$

III. CONDITION FOR LINK ADDITION AND DELETION

Link addition between an agent $i$ and an agent $j$ occurs only when the perceived utilities of both the agents increases, i.e.

$$\tilde{u}_{i}(g + \langle ij \rangle) > \tilde{u}_{i}(g) \quad (2)$$
$$\text{and}$$
$$\tilde{u}_{j}(g + \langle ij \rangle) > \tilde{u}_{j}(g). \quad (3)$$

For all agents $k$, except $i$ and $j$, the neighbourhood size ($n_{k}$) remains constant so their utility is the same and cancels out. Simpilifying (2) and (3) with (1), we get that link addition happens when

$$f_{ii}\lambda_{n_{i}} + f_{ij}\lambda_{n_{j}} > \frac{(f_{ii} + f_{ij}) * c * \lambda}{(1 - \lambda) * \beta} \quad (4)$$
$$f_{ii}\lambda_{n_{i}} + f_{ij}\lambda_{n_{j}} > \frac{(f_{ii} + f_{ij}) * c * \lambda}{(1 - \lambda) * \beta}. \quad (5)$$

Similarly, link deletion happens when

$$\tilde{u}_{i}(g - \langle ij \rangle) > \tilde{u}_{i}(g)$$
$$\text{and}$$
$$\tilde{u}_{j}(g - \langle ij \rangle) > \tilde{u}_{j}(g).$$

Thus, equivalent to (4) and (5) we get link deletion conditions as below.

$$f_{ii}\lambda_{n_{i}} + f_{ij}\lambda_{n_{j}} < \frac{(f_{ii} + f_{ij}) * c * \lambda}{(1 - \lambda) * \beta} \quad (6)$$
$$f_{ii}\lambda_{n_{i}} + f_{ij}\lambda_{n_{j}} < \frac{(f_{ii} + f_{ij}) * c * \lambda}{(1 - \lambda) * \beta}. \quad (7)$$

IV. SUFFICIENCY CONDITION FOR LINK ADDITION

Theorem 1. For any two agents $i$ and $j$, let $t_{1} = \max (n_{i}, n_{j})$, $t_{2} = \min (n_{i}, n_{j})$ and $f_{ij} > 0$. If $t_{1} < \frac{\ln\left(\frac{c}{1 - \lambda_{i} \lambda_{j}}\right)}{\ln\lambda}$, the link addition conditions (4) and (5) will be true, and hence, a link will be formed between $i$ and $j$.

Proof.

Case 1: If $n_{i} > n_{j}$, then $t_{1} = n_{i}$ and $t_{2} = n_{j}$. Thus,

$$\lambda_{n_{i}} < \lambda_{n_{j}} \quad (\text{since } 0 < \lambda < 1)$$
$$\text{or}$$
$$f_{ii}\lambda_{n_{i}} < f_{ij}\lambda_{n_{j}} \quad (\text{since } f_{ii} > 0).$$

Also, $f_{ij}\lambda_{n_{i}} < f_{ij}\lambda_{n_{j}} \quad (\text{since } f_{ij} > 0)$. Combining above two we get

$$(f_{ii} + f_{ij})\lambda_{n_{i}} < f_{ii}\lambda_{n_{i}} + f_{ij}\lambda_{n_{j}} < (f_{ii} + f_{ij})\lambda_{n_{j}} \quad (8)$$
$$\text{and}$$
$$(f_{ii} + f_{ij})\lambda_{n_{i}} < f_{ii}\lambda_{n_{i}} + f_{ij}\lambda_{n_{j}} < (f_{ii} + f_{ij})\lambda_{n_{j}}. \quad (9)$$

Using (4)–(5) in (8)–(9) we get sufficiency condition for link addition as below.

$$(f_{ii} + f_{ij})\lambda_{n_{i}} > \frac{(f_{ii} + f_{ij}) * c}{(1 - \lambda) * \beta}. \quad (10)$$

On solving (10), we get that link between $i$, $j$ will happen if

$$n_{i} < \frac{\ln\left(\frac{c}{1 - \lambda_{i} \lambda_{j}}\right)}{\ln\lambda}.$$ 

Case 2: If $n_{i} \leq n_{j}$

Since (4) and (5) are symmetric, interchange $i$ and $j$.

$\square$

Similarly, we can derive sufficiency conditions when $f_{ij} < 0$ above as well as for link deletion. Algorithm 1 lists the steps for reaching a bilaterally stable network. That is, when no agent has any incentive to add or delete a link.

V. CASE STUDY

Consider five agents as follows: $a, b, c, d$, and $e$. Assume that social relationships between these agents is captured in the Social Range Matrix $F$ given in Table I. As an example, the way to read this matrix is as follows: $a$ is a friend of $b$ and $d$, while it is an enemy of $c$ and $e$.

Let us also assume that $c = 0.01, \lambda = 0.2, \epsilon = 0.1$, and $\beta = 0.1$. Initially all agents are isolated (i.e., the starting
Algorithm 1: Pseudo code to arrive at a bilaterally stable network.

Input : $c, \lambda, \beta, F$, starting network, flag = 1
Comment: $i$ and $j$ are agents, flag = 1 means network is not bilaterally stable

while flag == 1 do
  flag = 0
  for $i = 1$ to $n$ do
    for $j = 1$ to $n$ do
      if $i \neq j$ then
        if link is absent between $< i, j >$ then
          Check link addition conditions (4) and (5) for $i, j$ and add link if they are true.
          flag = 1
        end
      end
      if link is present between $< i, j >$ then
        Check link deletion conditions (6) and (7) for $i, j$ and delete link if they are true.
        flag = 1
      end
    end
  end
end

TABLE I: Social Range Matrix $F$

|   | a | b | c | d | e |
|---|---|---|---|---|---|
| a | 1 | -1| -1| -1| -1|
| b | -1| 1 | -1| -1| -1|
| c | -1| -1| 1 | -1| -1|
| d | -1| -1| -1| 1 | -1|
| e | -1| -1| -1| -1| 1 |

network is a null network). We follow the procedure described in Algorithm 1 and obtain the bilaterally stable network $g'$ shown in Fig. 1. With the same set of parameters and $F$, by starting from a random non-null network, we obtain a different bilaterally stable network than in Fig. 1. This implies that bilaterally stable networks are not necessarily unique.

By looking at the ratio of $\frac{\lambda}{\beta}$, where $0 < c, \beta < 1$, $\lambda = 0.2$, and above $F$, we found that for $0.044 < \frac{\lambda}{\beta} < 0.089$, Algorithm 1 runs into infinite loop implying that no bilaterally stable network is possible. Interestingly, for all other values of $\frac{\lambda}{\beta}$ we found that at least one bilaterally stable network exists. If $c > \beta$, $\lambda = 0.2$, and above $F$, then we get a unique bilaterally stable network in which all pair of agents are enemies.

VI. UNIQUE BILATERALLY STABLE NETWORKS

Lemma 1. If $\frac{\lambda}{\beta} < (1 - \lambda)^{N - 2}$ and if all agents are friends of each other (i.e. $f_{ij} = 1$), then a complete network is the unique bilaterally stable network.

Proof. First, consider the link addition conditions stated in (4) and (5). We can observe that, as $n_i$ and $n_j$ increases, the L.H.S of (4) and (5) decreases. This implies that agents have an incentive to increase their neighborhood size.

If the link addition conditions (4) and (5) are true for $n_i = n_j = N - 2$ (both agents have the neighborhood size one less than the maximum possible size[4] then they will be true for all values of $n_i, n_j < N - 2$ (if the smaller value of the L.H.S is greater than the R.H.S., bigger values will be greater).

Solving (4) – (5) with $n_i = n_j = N - 2$, i.e.

$$f_{ii}(\lambda^{N - 2} + f_{ij}\lambda^{N - 2}) > \frac{(f_{ii} + f_{ij})\cdot c}{(1 - \lambda) \cdot \beta}$$

gives

$$\frac{c}{\beta} < (1 - \lambda)\lambda^{N - 2}.$$  

A similar analysis can be done with the deletion conditions.

Lemma 2. If $\frac{\lambda}{\beta} > (1 - \lambda)$ and if all agents are friends of each other (i.e. $f_{ij} = 1$), then the empty network is the unique bilaterally stable network.

Proof. Similar to Lemma 1.

Corollary 1. If $\frac{\lambda}{\beta} > \frac{1}{1 - f_{ij}}(1 - \lambda)$ and if $f_{ij}$ is either 1 or -1, then the bilaterally stable network will be the one where all the pairs of enemies form links.

VII. FUTURE WORK

In this work, we have extended the social storage model proposed in [4] to include heterogeneous behavior (variety of social relationships). After that, we have analyzed this model using their solution concept of bilateral stability. The preliminary results here give tremendous insight into how a endogenously built social storage system would emerge.

Future work involves further analyzing the stability of such networks as well as studying contentment (when everyone has maximized their utility) and efficiency (where the total utility of the network is maximized) in this new heterogeneous agent scenario.

REFERENCES

[1] N. Tran, F. Chiang, and J. Li, “Efficient cooperative backup with decentralized trust management,” ACM Transactions on Storage, vol. 8, no. 3, pp. 8:1–8:25, 2012.

[2] R. Gracia-Tinedo, M. Sánchez-Artigas, A. Moreno-Martínez, and P. García-Lopez, “Friendbox: A hybrid F2P personal storage application,” in 12th IEEE International Conference on Peer-to-Peer Computing (P2P), IEEE, 2012, pp. 131–138.

[3] R. Sharma, A. Datta, M. DeH’Amico, and P. Michiardi, “An empirical study of availability in friend-to-friend storage systems,” in IEEE International Conference on Peer-to-Peer Computing (P2P), 2011, pp. 348–351.

[4] P. Mane, K. Ahuja, and N. Krishnamurthy, “Unique stability point in social storage networks,” 2017, 33 Pages. [Online]. Available: [arxiv.org/abs/1703.07689](https://arxiv.org/abs/1703.07689)

[5] P. Kuznetsov and S. Schmidt, “Towards network games with social preferences,” in 17th International Colloquium on Structural Information and Communication Complexity (SIROCCO), B. Patt-Shamir and T. Ekim, Eds. Springer: Berlin Heidelberg, 2010, pp. 14–28.

1Neighborhood size of $N - 1$ is a complete network, and hence, is not included in the analysis here.