GAME THEORETICAL ANALYSIS OF RESOURCE ALLOCATION IN THE INTERPLANETARY FILE SYSTEM

by

David M. Grisham

A thesis submitted to the Faculty and the Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirements for the degree of Master of Science (Computer Science).

Golden, Colorado

Date __________________________

Signed: __________________________

David M. Grisham

Signed: __________________________

Dejun Yang

Thesis Advisor

Golden, Colorado

Date __________________________

Signed: __________________________

Tracy Camp

Head

Department of Computer Science

ABSTRACT

Welcome to my paper!

TABLE OF CONTENTS

ABSTRACT iii

LIST OF FIGURES 1

CHAPTER 1 INTRODUCTION 2

1.1 IPFS 2

1.2 Bitswap 3

1.3 Objectives 3

CHAPTER 2 SYSTEM MODEL 5

2.1 Network Graph 5

2.2 Reputation 5

2.3 Game Formulation 6

LIST OF FIGURES

Figure 1.1 The IPFS Stack 3

CHAPTER 1

INTRODUCTION

The Internet is perhaps the largest and most consistent network that has ever existed. Unfortunately, it predominantly runs on the outdated hypermedia distribution protocol HTTP. The goal of the InterPlanetary File System (IPFS) is to upgrade the Internet to a distributed peer-to-peer system, thereby making it more robust and permanent. This new Internet would be a network of peers, as opposed to clients and servers, all sharing data between one another. In order for such a system to thrive, users must be cooperative and willing to share data with their peers. The goal of this project is to analyze the resource allocation options of peers interacting in an IPFS network. A combination of analytical and empirical methods will be used to glean insights into the generally intractable allocation decisions that users are presented with when participating in an IPFS network.

1.1 IPFS

IPFS is a peer-to-peer hypermedia distribution protocol developed by Protocol Labs. It is a content-addressed, versioned filesystem. While a variety of use cases exist for such a protocol, the most ambitious goal of the project is to replace HTTP as the primary file exchange protocol used in the Internet. This could ultimately result in the decentralization of the Internet.

IPFS synthesizes various technologies developed since the Internet's inception. These technologies include Git, BitTorrent, distributed hash tables (e.g. Kademlia), and self-certified filesystems. The IPFS stack is shown in Figure Figure 1.1. One way to conceptualize an IPFS network is as a Git repository shared within a torrent-esque swarm.

Figure 1.1 **The IPFS Stack** – Bitswap is at the exchange layer.

1.2 Bitswap

Bitswap is the block exchange protocol of IPFS. The most direct inspiration of this submodule is the BitTorrent peer-to-peer file distribution protocol. Bitswap is the layer of IPFS that incentivizes users to share data. A Bitswap *reciprocation function* determines which peers to send data to, and in what relative quantities. The input to the reputation function is a set of metrics that may be used to weigh peers – e.g. peer bandwidth, reputation, and/or location. The output is a set of weights, one for each peer, that assign the relative resource allocations for the peers. These weights are periodically recalculated to reflect changes in both the network and peer behavior.

1.3 Objectives

TODO: This section will likely need to be updated/replaced, as it was copied directly from the proposal. At the very least, I think it might need to be changed to be past-tense instead of future

For this project, I will take the initial steps toward understanding the behavior of users in an IPFS network as predicted by game theoretical models. This will involve a combination of analytical and empirical analyses, along with implementation of these ideas in the IPFS software. The analytical work will focus on repeated games and, potentially, evolutionary games, while the empirical work will take a simulation-based approach. I intend to use these methods to classify various Bitswap reciprocation functions and determine useful allocation behavior under certain conditions.

CHAPTER 2

SYSTEM MODEL

This section details the model currently used to describe Bitswap in this work. §2.1 models the IPFS network as a graph; §2.2 mathematically describes the peer-wise reputations and user interactions; and §2.3 formulates the problem as a game.

2.1 Network Graph

We model an IPFS swarm as a network $\mathcal{N}$ of $|\mathcal{N}|$ users. The graphical representation consists of

- *nodes* representing users, and
- *unweighted, undirected edges*, each of which represents a peering between two users.

A user's *neighborhood* is their set of peers, i.e. the set of nodes that the user is connected to by an edge. User i's neighborhood is denoted by $\mathcal{N}_i$, where $\mathcal{N}_i \subseteq \mathcal{N}$.

2.2 Reputation

We break Bitswap interactions into discrete rounds, with a single round denoted by a nonnegative integer t. The following two points describe the way data distribution takes place in this Bitswap model. Each of these points simplifies the problem from the real-world scenario.

1. Each user j distributes exactly B_j bits, where $B_j > 0$, to each of their peers in a given round (and has sufficient resources to do so).
2. All users always have unique data that all of their peers want. So, when a user allocates b bits to a particular peer, that user has at least b bits that the peer wants.

We define b_{ij}^t as the total number of bits sent from user i to peer j from round 0 to $t-1$. User i maintains a Bitswap *ledger* l_{ij}^t, for each of its peers $j \in \mathcal{N}_i$, that stores the amount of data exchanged in both directions, i.e. $l_{ij}^t = (b_{ij}^t, b_{ji}^t)$.

Now we define the *debt ratio* d_{ji} from user i to peer j as

$$d_{ji}^t = \frac{b_{ij}^{t-1}}{b_{ji}^{t-1} + 1}$$

d_{ji}^t can be thought of as peer i's reputation from the perspective of user j. This reputation is then considered by user j's *reputation function* $S_j(d_{ji}^t, \mathbf{d}_j^{-i,t}) \in \{0, 1\}$, where $\mathbf{d}_j^{-i,t} = (d_{jk}^t \mid \forall k \in \mathcal{N}_j, k \neq i)$ is the vector of debt ratios for all of user j's peers in round t *excluding* peer i. The reputation function considers the relative reputation of peer i to the rest of j's peers, and returns a weight for peer i. This weight is used to determine what proportion of j's resources to allocate to peer i in round t. This means that the following must hold:

$$\sum_{i \in \mathcal{N}_j} S_j(d_{ji}^t, \mathbf{d}_j^{-i,t}) = 1$$

In other words, the sum of the weights that j assigns to its peers must be 1 (otherwise, j would allocate more than its available bandwidth to its peers).

Given all of this, we can calculate b_{ji}^{t+1} given b_{ji}^t, d_{ji}^t, and $\mathbf{d}_j^{-i,t}$:

$$b_{ji}^{t+1} = b_{ji}^t + S_j(d_{ji}^t, \mathbf{d}_j^{-i,t}) \times B_j$$

2.3 Game Formulation

The game model presented here is the product of multiple iterations that approached a balance between the accuracy of the model to the problem and model complexity. Previous modeling approaches included tools from evolutionary game theory and statistical mechanics on the high complexity end, and repeated games on the low accuracy end. While the current model uses a repeated game model as well, the strategy space has been modified to better fit the Bitswap scenario.[1]

All users in the network participate in the Bitswap game. The players in this game are the users in the IPFS network, and each player's strategy is the reciprocation function they choose to assign weights to their peers. The strategy space, then, is the space containing all pure functions of the form $S_j(d_{ji}^t, \mathbf{d}_j^{-i,t}) \in \{0, 1\}$. The Bitswap game is an *infinitely repeated*, *incomplete information* game in which users exchange data. One game takes place between each pair of peers for each round t. The *players* are the IPFS users in the network $\mathcal{N}$. The utility of player i in round t is the sum of all of the data that i is sent by its peers in that round:

$$u_i^t = \sum_{j \in \mathcal{N}_i} (b_{ji}^t - b_{ji}^{t-1})$$

The total utility of a peer over the entire repeated game, then, can be expressed in the following ways:

$$U_i = \sum_{t=0}^{\infty} u_i^t$$
$$U_i = \sum_{j \in \mathcal{N}_i} b_{ji}^{\infty}$$

The latter representation can be most directly thought of as the total amount of data that peer i receives from all of its peers over the entire duration of the Bitswap game.

[1] A description of the model from the previous iteration can be found at https://github.com/dgrisham/masters/tree/master/deprecated/analysis.