A Survey of Incentive Mechanisms in Peer-to-Peer Systems

Technical Report CS-2009-22

Muntasir Raihan Rahman
Cheriton School of Computer Science, University of Waterloo
E-mail: mr2rahman@cs.uwaterloo.ca

***Abstract*— The fundamental assumption that peer-to-peer (P2P) networks can thrive on voluntary contribution of altruistic peers can no longer be supported without considering the impact of rational behavior on such decentralized systems. This paper attempts to shed light on the impact of rational free-riding behavior of participating peers on the stability and existence of real-world peer-to-peer networks and the various attempts to cope with this problem. In particular, we focus on the economic principles that drive these problems, the various incentive mechanisms proposed to thwart these problems and analytical tools used to describe these rational manipulations in P2P systems.**

I. Introduction

The decentralized peer-to-peer model (P2P) of communication has emerged as a viable alternative to the traditional client-server based centralized model in the realm of large scale distributed systems. In a P2P system, a number of autonomous nodes form a self-organizing and self-sufficient networked system without any centralized authority. Due to this, the performance of such a decentralized system largely depends on the level of voluntary cooperation from the autonomous nodes [18], [10]. Traditional systems normally assume obedient users - users who abide by the governing protocol and rules without considering their own utility. However this assumption seems unrealistic in P2P network settings where the peers interact with one another in diverse situations and with different levels of collaboration and competition. As a result researchers were forced to abandon the concept of obedient users and turn their attention to *rational users* who attempt to maximize their own utility by deviating from the standard protocols, thereby reducing the overall social welfare of the system. To this end, it seems that in a rational setting, individual rationality is in direct opposition with social welfare.

Users who try to benefit from a system without making any contribution to the system are termed as *free-riders*. Various measurement studies have confirmed large scale free-riding behavior in commercial P2P systems that do not take into account rational behavior [2]. The prevailing problem of free-riding in P2P networks and its related counterparts (including white-washing and sybil attacks) is the main focus of this paper. We attempt to identify the reasons of this prevalent behavior and how current systems attempt to cope with it and we also touch upon some analytical models that have been proposed in the literature to describe these rational behaviors in P2P systems.

The rest of the paper is organized as follows. In Section II, we describe the free-riding problem in detail, including the related problems of white-washing and sybil attacks. Following that, in Section III we provide a high level classification of the various incentive mechanisms that have been proposed in the literature. Then in Section IV, we give detailed exposition of four real P2P systems and explain their basic architecture, design tradeoffs and their resilience against different rational attacks. Following this, in Section V, we discuss a number of analytical tools, mainly borrowed from micro-economics [1] and game theory [18] that have been used to explain these rational behaviors in P2P systems. We also include a comprehensive comparison of different incentive mechanisms and real systems with respect to their tolerance against the mentioned rational attacks in Section VI. We conclude in Section VII.

II. Problem Description

A. Free Riding

The public world was introduced to the realm of large scale peer-to-peer networking with the advent of Napster and Gnutella. Within a short time of its release, the Gnutella network found about 75% of its users free-riding, i. e. , downloading files from the network without uploading any content [2]. In game theoretic [18] terms, it is a *dominant strategy* for users to download files without making any contribution by uploading files. As a result all individuals can reason in this way and and free ride on the contributions of other *benevolent users*, ultimately resulting in overall system performance degradation and making everyone a loser, a situation referred to as the *tragedy of the digital commons* [14]. The free-riding problem is not just observed in P2P systems. This type of rational behavior also appear in the context of strategic network formation, selfish routing, mobile ad-hoc networks and congestion control. However the prevalence of free-riding behavior in P2P systems can be attributed to some specific characteristics, including decentralization, high churn rate (users dynamically joining and leaving the system), availability of cheap identities (pseudonyms), hidden actions (e.g. nodes not forwarding messages to reduce computing and communication cost), and collusion (users forming groups to maximize benefit). Due to the rational behavior of strategic users, it was quickly realized that some sort of incentive mechanism was required to over come the free-riding problem

and to encourage users to cooperate. As a result the next generation of P2P file sharing networks incorporated incentive mechanisms based on currency and reputation. For example, in *Mojonation* peers earn currency by making contributions and use the earned currency to purchase service from other peers. On the other hand in *KaZaA*, peers increase their reputation by uploading and later use their high reputation scores during downloading. The *BitTorrent* [7] system goes beyond these simple mechanisms and proposes a novel incentive mechanism based on the TFT (Tit For Tat) strategy which is modeled analytically using the Iterated Prisoner's Dilemma (IPD) Game (V-A.1).

B. Availability of Cheap Pseudonyms: White Washing

In a whitewashing attack, a free-rider continuously leaves and rejoins the P2P network under new identities in order to avoid the penalties imposed on free-riders [12]. Reputation systems, which are based on indirect reciprocity are particularly vulnerable to these kind of attacks, since a free-rider can clean her bad reputation by leaving the system and rejoining. The feasibility of the white-washing attack is attributed to the availability of low cost identities or cheap pseudonyms [12]. Normally there are two ways to counter the white-washing attack. The first approach is to necessitate the use of free but irreplaceable pseudonyms via the assignment of strong identities by a centrally trusted authority, however this reduces the decentralized nature of P2P systems and introduces a single point of failure. Without a trusted third party, the only option left is to impose penalties on all newcomers, which will include white-washers. This might seem harsh and detrimental to the scalability of the system and it has been shown to increase the total social cost of the system [12]. In [9], the authors have developed a simple economic model of user behavior in order to explain the prevalence of free-riding and white-washing in P2P systems. Their work elicits two major insights, first, when the level of generosity in the system is low, a mechanism that penalizes free-riders can greatly improve system performance and second, if identities are free, penalizing all newcomers blindly can effectively discourage white-washers and will incur increases in social cost for only limited and rare scenarios and high churn rates.

C. Sybil Attack

The *Sybil Attack* was first introduced by John R. Douceur in [8]. In a sybil attack, an attacker thwarts the reputation and sharing mechanism of a P2P network by creating a large number of identities or pseudonyms, using them to gain a disproportionately large influence. The level of damage of this attack depends on the availability of cheap pseudonyms in a system [12] (see II-B for more details). In [8], it is argued that without a trusted centralized authority, Sybil attacks are always possible except under the unrealistic assumptions of huge resources and coordination among entities. This is proved by a series of four simple lemmas, which we state below [8].

Lemma 1: If ρ is the ratio of the resources of a malicious entity f to the resources of a minimally capable entity, then f can present $g = \lfloor \rho \rfloor$ distinct identities to a local entity l.

Lemma 2: If a local entity l accepts entities that are not validated simultaneously, then a single malicious entity f can present an arbitrary large number of distinct identities to entity l.

Lemma 3: If a local entity l accepts any identity vouched for by q accepted identities, then a set F of faulty (malicious) entities can present an arbitrarily large number of distinct identities to l if either $|F| \geq q$ or the collective resources available to F at least equal those of $q+|F|$ minimally capable entities.

Lemma 4: If the correct entities in a set C do not coordinate time intervals during which they accept identities, and if local entity l accepts any identity vouched for by q accepted identities, then even a minimally capable faulty entity f can present $g = \lfloor |C|/q \rfloor$ distinct identities to l.

Although these negative results may cast a dark shadow on system designers who rely on altruistic behavior from users, it is not the end of the story. A local entity's ability to thwart a sybil attack can be enlarged with increased resources. For example, an entity can issue resource demanding challenges to validate identities, like challenging users to solve a unique computational puzzle. Recently Yu et. al. have proposed a novel protocol *SybilGuard* for limiting the corrupting influence of sybil attacks [27], based on social networks among user identities, where an edge between two identities indicate a trusted relationship. The success of this technique depends on the existence of small cuts in the relationship graph between malicious sybil nodes and honest nodes, since a malicious node can create many identities (nodes in the graph), but few trust relationships (edges in the graph). Sybilguard uses this graph-theoretic property to bound the number of identities a malicious node can create. The notion of sybil-proofness, that is, robustness against sybil attacks has recently been formalized in [6], where the authors argue that if reputation is solely determined by the graph structure and edge costs, then there is no mechanism that can thwart sybil attacks. On the other hand, if reputations are computed with respect to a fixed node in the graph, then sybil-proofness is guaranteed subject to several constraints. For example, the authors in [11] have devised some network flow-based algorithms for trust propagation that satisfy the conditions of sybil-proofness.

III. Incentives for Cooperation

This classification scheme has been adapted from [10].

A. Inherent Generosity

The *Warm-Glow Model* was proposed by Andreoni [3] to explain why some users gain altruistic utility from the mere act of giving. Based on this model, the authors in [9] developed a mathematical model to explain the free-riding and white-washing behavior in P2P systems, taking user generosity into account. The model is capable to analytically determine the percentage of free-riders in a system based on the probabilistic population distribution. The main insight that was elicited from that model is that if the societal generosity level is below a certain threshold, then the system ceases to exist as a useful artifact since the number of selfish users increase unboundedly.

On the other hand above that threshold, the performance of the system increases with higher levels of generosity and follows the well known principle of diminishing marginal returns [1].

B. Monetary Payment Schemes

In monetary payment schemes, users are required to pay in some form of virtual currency to get specific services from other peers. Monetary schemes allow for rich economic mechanisms based on accounting and micro-economic infrastructures but suffer from scalability. Much of the work in this line assume feasibility of micro-payment schemes without implementation details. Other problems include the hidden costs of service providers and how to motivate them to reveal their true costs (via mechanism design, see section V-B). The system Karma [25] falls within this category and a detailed explanation of the system is provided in Section IV-D.

C. Reciprocity-Based Schemes

In reciprocity-based schemes, peers maintain behavior histories of other peers and utilize this information for decision making processes. These schemes can be based on direct reciprocity or indirect reciprocity. In direct reciprocity, user X decides on the cooperation level with Y based only on the service received from Y in the past. On the other hand indirect reciprocity schemes also take into account the service Y has provided to other users of the system apart from X.

1) Direct Reciprocity: Direct reciprocity schemes are mainly applicable for applications that sustain for long durations since they provide adequate opportunities for reciprocation between two users. BitTorrent falls within this category (see section IV-B). Although experimental and analytical studies have found increased level of cooperation in these systems, a recent study demonstrates that a free-rider can still accomplish download completion times comparable to a regular altruistic contributor [15].

2) Indirect Reciprocity: Many indirect reciprocity schemes have been proposed in the literature and we describe in detail one such system in the next section IV-A. These systems are often called reputation systems and they differ from their direct counterpart in the computation of reputation scores and mapping of scores to strategies. Indirect reciprocity systems are more scalable than direct reciprocity systems, however they rely on third party observations and must handle trust issues which are absent in the direct reciprocity systems. The main disadvantage of these systems are that they are extremely vulnerable to whitewash attacks (Section II-B)and sybil attacks (Section II-C).

IV. Detailed Exposition of some Incentive Aware P2P Systems

This section discusses in detail four P2P systems that explicitly employ some of the incentive mechanisms discussed in the previous section.

A. Credence

Credence [26] is a P2P distributed reputation system, designed to thwart content pollution in P2P file sharing systems. Credence allows a peer to evaluate the authenticity of a file or any other online content based on the accuracy of the purported description of the object in contrast to the object itself. Members of the Credence Network vote on objects, the system then collates these votes and weights them via a similarity measure that weighs votes from like minded peers highly, whereas votes from vote-spammers and other intruders are weighed poorly. This new vote correlation scheme provides users a strong incentive to vote honestly and to mitigate the negative impact of malicious users. The authors have implemented Credence on top of the LimeWire client for the Gnutella network. The implemented client provides a peer assisted judgment mechanism that ensures whether an online object possesses desired authenticity properties and empowers users to evaluate object search results before actually downloading them.

The authors roughly define pollution as any file with content that does not match its label published description. On the other hand an authentic file has content that accurately matches its metadata description. One key observation of the authors is that pollution in current P2P file sharing networks can be easily detected by honest users without any sophisticated ranking or mathematical technique. The Credence system relies on the individual users as the first line of defense against file pollution attacks. After a user downloads a file, she is given a single chance to submit a vote to the Credence system: a positive thumbs-up for an authentic file, and a negative thumbs down for a corrupt or polluted file. Votes are cryptographically signed to ensure non-repudiation and to prevent sybil attacks [8]. Credence uses these signed votes to determine the authenticity of a file and displays a rating for each file based on its rating. The client software executes a search for votes and randomly downloads a number of votes and finally aggregates all these votes for a unified measure of authenticity for an online object. To guard against malicious or byzantine faulty peers, each peer is assigned a correlation coefficient, reflecting the historical usefulness of the peer's vote via indirect reciprocity. This scheme discourages an attacker to lie about the authenticity of the file. Due to the scalability and high churn rate of P2P networks, peers actually end up sharing and voting on few files over their short lifetime. This can pose trouble to a client willing to measure the trustworthiness of a file. To alleviate this problem, the authors have proposed to to use a technique called transitive correlation to rapidly disseminate information among small groups and help clients acquire historical information on a much larger scale. In Credence, a client repeatedly requests historical data from randomly selected peers which contains information about the peers past voting history and the relation of the peers with its neighbors. These data are then authenticated by the client and incorporated into the clients local database. This allows the client to leverage both the work done by other peers in evaluating files and the past behavior of the peers, without direct user interaction or complicated trust computations.

The authors of Credence have studied the impact of different rational attacks on the system. For example, a consistently lying peer is no more effective in the rational sense, than an honest peer, because the lying peer's votes are multiplied by a negative weight. Randomly generating multiple votes and launching sybil attacks will eventually lead to votes being discarded, since the peer correlation value tends to zero in these cases. As a result any rational attacker has strong incentive to vote honestly in the system. The authors have evaluated the effect of whitewashing attacks, which in this context means that an attacker votes correctly on a large set of files before endorsing a small set of invalid objects to pollute the system. However, it has been observed that the damage caused by the whitewasher is partially offset by the large number of correct votes that are required to launch the attack in the first place. Moreover multiple independent whitewashing attacks tend to annihilate each other.

B. BitTorrent

BitTorrent [7] is a distributed file downloading system where peers can download a file from each other in excess of the original source. The basic idea is that a server divides a file into discrete pieces and gives each piece of the file to a peer, which itself can supply that piece to other peers. Peers build up a file by requesting for missing parts from other peers. This simple scheme avoids a single point of failure and distributes the responsibility of providing a file among multiple peers, forcing them to cooperate to achieve a common goal. In other words the system designer achieves the goal of providing incentive to users to link the opposite activities of downloading and uploading file pieces. The basic incentive mechanism in BitTorrent is that a user's download bandwidth is proportional to her upload performance, which is an example of direct reciprocity modeled according to the Tit for Tat (TFT) strategy or the game theoretic Iterated Prisoners Dilemma (IPD) model [4]. To this day, BitTorrent is the most popular file downloading tool available and has been widely studied by academicians via theoretical and experimental analysis. There are four main reasons for this: first, the system has been designed with the assumption of rational profit maximizing users, second, the system designer has proposed a default strategy, i. e. a default BitTorrent client is available for download, third, the system is simple enough to be amenable to analytical, experimental and economic analysis [22], [13] and last but not least, the system is very popular among internet users.

The basic working principle of of BitTorrent is described briefly as follows: (for details see [7]). A system wide trusted tracker node maintains a list of peers that are active w.r.t to a given file. Completed peers and uncompleted peers are both tracked by the tracker. A new peer announces its arrival by contacting the tracker and then requests a random list of peers to download from. It then attempts to establish bi-directional TCP connections with the selected peers. Among these connections, a small subset of peers are internally marked as *unchocked*. An unchoked connection is one that where the other end point peer can also request a piece of a file and this peer will fulfil that request. An active peer that has completed downloading the file is called *altruistic* and will send pieces of the file to other incomplete peers based on their bandwidth. An incomplete peer on the other hand will rationally unchoke peers that provide the highest throughput. Both complete and incomplete peers use the optimistic unchoking strategy at a fixed interval to discover potentially better trading peers. This optimistic unchoking strategy forms the basis of the evolution of cooperation since by this mechanism a peer can form a direct reciprocity relation with an unknown peer and hope for cooperation in the future in the flavor of the Iterated Prisoners Dilemma game strategy. On a separate fixed interval, each peer keeps the k fastest peers unchoked, and chokes the remaining peers. BitTorrent has been shown to be vulnerable to different types of rational attacks including whitewashing attacks and sybil attacks.

C. BitTyrant

To-date BitTorrent [7] has been the most successful file distribution tool with explicit incentive mechanisms (Tit for Tat reciprocity strategy) to provide incentives to users to contribute authentic resources to the system. However a strategically designed client can thwart the BitTorrent incentive mechanism, as has been demonstrated in [21]. The authors in [21] exploited the fact that although the tit-for-tat (TFT) strategy provides fairness by balancing resource contribution with resource consumption and is an evolutionary stable strategy (ESS) [4], in practice high capacity users end up contributing more than they receive. This suggests that in BitTorrent, TFT does not perform as intended and might be exploited by rational users to improve performance. To this end, the authors in [21] have designed BitTyrant, a selfish client that attempts to demonstrate that in-fact incentives don't build robustness in bit-torrent. The main idea is that BitTyrant dynamically chooses the number and type of peers to send data, in contrast to the static approach in BitTorrent, where clients send data to a fixed number of clients in each TFT round, regardless of upload capacity. This dynamic adjustment algorithm maintains two statistical estimates, d, the rate at which peers provide data, and, u, the rate required to earn reciprocity. The highest capacity peers are then selected based on these estimates and data is send to them at the minimum rate that will lure them to reciprocate. At the end of each round, if a peer does not reciprocate, her u value is increased and if a peer un-chokes the requesting client, then her u value is decreased. The authors have compared the relative performance of BitTorrent and BitTyrant on more than 1000 real-world swarms as well as on PlanetLab [19] based synthetic swarms. The results confirmed that BitTyrant shows dramatic performance improvement over BitTorrent (around 70%) and it was observed that some clients finished downloads about 3 times faster. The authors have elicited some key insights that are responsible for the increased performance, e.g. (a) BitTyrant provides consistent performance in the long run, (b) it can identify the point of diminishing marginal returns for high capacity clients, and (c) low capacity peers can also benefit from BitTyrant. However the increased performance comes at a cost. For example new users can experience lengthy bootstrapping periods, and the peering relationships can be

unstable over the long run. However like BitTorrent, BitTyrant is vulnerable to white-washers and recent studies have shown it to be vulnerable to sybil attacks as well.

D. Karma

A P2P system can enforce rational users to contribute resources to the global resource pool by employing a economic currency scheme. Like the traditional prevalent economic system, peers in this type of system earn virtual currency by contributing resources to the system and spend the currency to purchase resources from the system. KARMA [25] proposes a general economic framework for combating free-riders in p2p systems by keeping track of resource contribution and resource consumption of each member of the system. This is achieved by representing the overall performance of each participant via a single metric called *karma*. However the karma values for each node are maintained by a a set of other nodes (called the bank-set) who are collectively responsible for continuously increasing and decreasing the karma value for that node as it contributes and consumes resources. Initially a user is awarded a seed amount of karma when she joins the system (this can encourage white washing). A user is not allowed to purchase via an (atomic) transaction if she does not have enough karma to pay for the resource which forces the participants to maintain a fine balance between their resource contribution and consumption. The main design issues of KARMA include the absence of any centralized trusted third party since the banking functionality is performed by other members, the prevalence of replication for fault tolerance and security against tampering of karma values by bank-set members. The KARMA design assumes that there are at least k nodes in the system at all time instances and that a certain fraction of these nodes ar non-malicious. The bank-set information is maintained via a Distributed Hash Table (DHT) (e.g. chord [24]) data structure where each node is mapped to bank-sets. The k closest nodes in the identifier space of each node A constitute the bank-set for A. Each member of the bank-set of A stores the current value of A's karma signed with A's private key and a transaction log of A's recent dealings. It also stores the current epoch number, which is a fixed time span agreed upon before hand. Currency adjustments are made at the end of each epoch in order to cope with inflation and deflation which can occur when nodes use up their karma and leave the system or accrue karma and leave. The adjustment is made by applying a correction factor at the end of each epoch. KARMA maintains file information using a *fileId* for each file. When a node joins it associates its id with the fileId's of all files that it possesses. A node willing to download a file acquires a list of potential up-loaders and initiates a lowest bid auction or a second price auction (Vickrey auction) [18] to select the peer to download from. Once a peer has been selected, the Karma file exchange protocol is started which tolerates temporary debit/credit inconsistencies during the exchange and avoids complicated Byzantine consensus protocols. Basically the protocol works as follows: the karma transfer from A to B is initiated when A sends B a signed message authorizing $Bank_A$ to transfer a given amount of karma to B, which forwards this message to $Bank_B$ which in turn contacts $Bank_A$. If A has sufficient karma in its account, the amount is deducted from A's account and credited to B's account. For security all the messages are authenticated to avoid repudiation and other security risks.

The main advantage of Karma is that by keeping track of virtual currency, it forces peers to maintain a check and balance between its resource contribution and usage. However there is an overhead since peers are also required to act as bankers for other nodes and from a game theoretic perspective, no peer has any incentive to take this additional responsibility. As a result Karma may introduce additional free-riding behavior in an attempt to cope with free-riding itself!

Karma has been argued to be resilient against a number of potential attacks, including replay attacks, malicious providers and consumers, corrupt bankers in the bank-set and denial of service attacks. Karma permits Sybil attacks on a limited scale which could be a potential drawback of the system. To date there has been no work on identifying the system's resilience against other important rational attacks including whitewashing and hidden actions which could be a possible avenue for future research.

V. Analytical Tools

In this section we discuss some of the prevalent analytical tools that have been used in the literature to analyze the properties and performance of incentive mechanisms for P2P systems.

A. Game Theory

Without any doubt, game theory is the most comprehensive analytical tool available for the study of incentive mechanisms in P2P systems. The dominant game-theoretic model used in this arena is the Prisoner's dilemma (PD) model, the iterated or repeated prisoner's dilemma (IPD) model (also known as Tit-for-Tat) and and its many variants. However the question arises as to whether these are the right models or whether there are other models that need to be investigated. A recent approach proposed by Peterson et. al. [20] points out some of the limitations of the Tit-for-Tat (TFT) strategy and advocates a new approach seeking the globally optimal outcome known as the *common good*.

1) Prisoner's Dilemma and Tit-for-Tat: The most widely studied game theoretic model for P2P incentives is the Two Person Prisoner's Dilemma game. The story behind the game is that two prisoners are on trial and they have two choices: confessing the crime or remaining silent. If they both remain silent, then the authorities cannot prove anything and they both get nominal punishment of 1 year each. If only one of them confesses then her term is reduced to say 0 years but the other prisoner gets 5 years. Finally if they both confess then they both get 3 years. The payoff matrix of the game is given in table I.

As can be seen from the payoff matrix, the dominant strategy for each player is to defect, since any player always gains more (or loses less) by defecting rather than cooperating regardless of the other players choice, because in the game

	Cooperate	Defect
Cooperate	$R = 3,\ R = 3$	$S = 0,\ T = 5$
Defect	$T = 5,\ S = 0$	$P = 1,\ P = 1$

TABLE I

PAYOFF MATRIX FOR THE PRISONER'S DILEMMA GAME [15]. R FOR REWARD, T FOR TEMPTATION TO DEFECT, S FOR SUCKER'S PAYOFF, AND P FOR PUNISHMENT FOR MUTUAL DEFECTION.

we have $T > R$ and $P > S$. This apparent impossibility of cooperation among rational peers can be overcome if the mutual interactions are repeated over time, and is modeled via a repeated game (Iterated Prisoner's Dilemma). To find whether this repeat interaction can generate cooperation among users, Alexrod [4] performed some online experiments by running computer tournaments in which pairs of players were subjected to iterated exchanges over time. After evaluating the results, it was seen that the long term winning strategy was the TIT-FOR-TAT (TFT) strategy, which can be stated as:

In the first interaction, always cooperate. After that do whatever the other player did in the previous interaction.

This deceptively simple strategy turns out to be evolutionary stable (ESS) and can out perform other rational strategies. In the P2P perspective, this strategy means that strangers are always treated benevolently. That is a peer always cooperates with a newcomer (e.g. by letting him download) and follows the newcomer's strategy in following interactions. However this strategy of cooperating with newcomers blindly can be exploited by white-washers (see section II-B).

The authors in [15] have attempted to model the BitTorrent incentive mechanism via the Iterated Prisoner's Dilemma game. Let $d > 0$ denote the download utility and $u > 0$ denote the cost of uploading. Then the payoff values will be: $R = d - u$, $T = d$, $S = -u$, $P = 0$. These value assignments meet the IPD constraints: that is, $T > R > P > S$ and $2R > S + T$, where the second condition is needed since otherwise two players would earn more by alternating between cooperation and defecting in consecutive rounds rather than always cooperating.

Various other game-theoretic models have been proposed in the literature [22], [5], however none of them are as popular as the Prisoner's Dilemma models. In [22], the authors present a simple fluid model to study the scalability, performance and efficiency of P2P file sharing systems like BitTorrent. In [5], the authors propose a differential service based incentive mechanism for P2P systems and prove that any P2P system with differential incentives will eventually reach a Nash equilibrium (a joint strategy from which no user has any incentive to unilaterally deviate from, provided that other users don't deviate) [16], [18].

B. Mechanism Design: Inverse Game Theory

The theory of mechanism design [18] provides an elegant mathematical framework for designing games, where the behavior of rational players result in the socially desirable outcome (i. e. the player's incentives are aligned with the mechanism designer's goal). This alignment is accomplished by forcing players to pay a price which is deducted from their utility. Classical mechanism design however ignores the computational complexity of mechanism design algorithms, which led computer scientists to develop the theory of algorithmic mechanism design (AMD) [17] which attempts to reconcile computer science and economics by developing a formal computational model that combines incentive compatibility (the notion that no player in a game can gain more utility by lying rather than declaring their true valuation of a service or desired object) and computational tractability. Shneidman et al. [23] have proposed to use mechanism design as a network design tool to help deal with rational nodes in P2P networks. To this end, they have identified various node types to help identify strategic behavior in systems. For example *correct/obedient* nodes are those that abide by the network protocol, whereas *faulty* nodes can either stop working or act arbitrarily (byzantine failure). However the node types that are more relevant from an economic perspective are the *Rational* nodes (attempting to maximize utility) and *irrational* nodes (nodes that behave strategically but do not follow the mechanism protocol). The authors have also identified some open problems in Distributed Algorithmic Mechanism Design (DAMD) (the distributed counterpart of AMD where the task of mechanism design is entrusted to the selfish agents themselves). Mechanism design can also be helpful in explaining the performance of monetary exchange schemes (see Section III-B) and direct reciprocity systems (see Section III-C.1).

VI. COMPARISON

In this section, we provide a comprehensive tabular comparison (tables II and III) of the attack resilience of the various proposed incentive mechanisms and the four real systems that we have described in a previous section (Section IV). To the best of our knowledge, this is the first such tabular comparison presented in the literature. This can be useful for identifying future research avenues and for pointing out the limitations of current systems. It should be noted that some of the entries in the table are marked as *Partially*, which denotes that the corresponding incentive mechanism or real system is not totally resilient against the particular attack, but rather it can only resist the attack on certain cases or under special conditions. For example, Karma (IV-D) permits sybil attacks (II-C) on a limited scale, which means it is only partially resistant to sybil attacks. An entry labeled *Unknown* signifies that the resilience of the system or incentive mechanism against the corresponding rational attack hasn't yet been addressed in the research literature and therefore is a strong candidate for future research directions.

VII. CONCLUSION

In this paper, we have attempted to provide a comprehensive survey of the prevalent free-riding problem in decentralized P2P networks and the various incentive mechanisms proposed for existing systems to cope with these problems. We have explained in detail the architecture and attack resilience capability of some well-known P2P networks, including the popular file-sharing system Bit-Torrent. We have also touched

Mechanism vs Attack	Rationality	Free Riding	White Washing	Sybil Attack
Inherent Generosity	No	No	No	No
Monetary Payment	Yes	Yes	Partially	Partially
Direct Reciprocity	Yes	Partially	Partially	Partially
Indirect Reciprocity	Yes	Partially	No	No

TABLE II
RESILIENCE OF INCENTIVE MECHANISMS AGAINST VARIOUS ATTACKS

System vs Attack	Rationality	Free Riding	White Washing	Sybil Attack
Credence	Yes	Unknown	Partially	Partially
BitTorrent	Yes	Partially	No	No
BitTyrant	Yes	Partially	No	No
KARMA	Yes	No	Unknown	Partially

TABLE III
TOLERANCE OF EXISTING SYSTEMS AGAINST VARIOUS ATTACKS

upon the various economic and mathematical models that have been put forward by researchers to explain these problems in P2P networks. We have also provided a comparison of different incentive mechanisms and real systems with respect to resilience against free-riding and other attacks. In summary, rational user behavior has the potential to disrupt any P2P system that relies on the assumption of altruistic and benevolent peers. Although there exists a large body of literature and real deployed systems in this research area, we believe that to date there is no robust incentive mechanism that strongly prohibits users to free-ride and encourages them to make contributions. The main analytical tools used so far are game-theoretic in nature and are borrowed from economics. It can be inferred that some other analytical models and experimental tools from related disciplines, e. g. sociology, finance and behavioral science might also be useful to shed light on the rational behaviors in these P2P systems. Finally, the main objective of this line of investigation is obvious: to discover robust and computationally tractable incentive mechanisms that force rational users to align their personal interest with the global system wide objective of social welfare maximization.

REFERENCES

[1] Mas-Colell A., Whinston M. D., and Green J. R. *Microeconomic Theory*. Oxford University Press, Oxford, UK, 1995.

[2] E. Adar and B. Huberman. Free riding on gnutella, 2000.

[3] James Andreoni. Giving with impure altruism: Applications to charity and ricardian equivalence. *The Journal of Political Economy*, 97(6):1447–1458, 1989.

[4] R. Axelrod. The evolution of cooperation. 1984.

[5] Chiranjeeb Buragohain, Divyakant Agrawal, and Subhash Suri. A game theoretic framework for incentives in p2p systems. In *P2P '03: Proceedings of the 3rd International Conference on Peer-to-Peer Computing*, page 48, Washington, DC, USA, 2003. IEEE Computer Society.

[6] Alice Cheng and Eric Friedman. Sybilproof reputation mechanisms. In *P2PECON '05: Proceedings of the 2005 ACM SIGCOMM workshop on Economics of peer-to-peer systems*, pages 128–132, New York, NY, USA, 2005. ACM.

[7] Bram Cohen. Incentives build robustness in bittorrent, 2003.

[8] John R. Douceur. The sybil attack. In *IPTPS '01: Revised Papers from the First International Workshop on Peer-to-Peer Systems*, pages 251–260, London, UK, 2002. Springer-Verlag.

[9] M. Feldman, C. Papadimitriou, J. Chuang, and I. Stoica. Free-riding and whitewashing in peer-to-peer systems. *Selected Areas in Communications, IEEE Journal on*, 24(5):1010–1019, May 2006.

[10] Michal Feldman and John Chuang. Overcoming free-riding behavior in peer-to-peer systems. *SIGecom Exch.*, 5(4):41–50, 2005.

[11] Michal Feldman, Kevin Lai, Ion Stoica, and John Chuang. Robust incentive techniques for peer-to-peer networks. In *EC '04: Proceedings of the 5th ACM conference on Electronic commerce*, pages 102–111, New York, NY, USA, 2004. ACM.

[12] Eric J. Friedman and Paul Resnick. The social cost of cheap pseudonyms. *Journal of Economics and Management Strategy 10(2)*, pages 173–199, August 2000.

[13] Lei Guo, Songqing Chen, Zhen Xiao, Enhua Tan, Xiaoning Ding, and Xiaodong Zhang. Measurements, analysis, and modeling of bittorrent-like systems. In *IMC'05: Proceedings of the Internet Measurement Conference 2005 on Internet Measurement Conference*, pages 4–4, Berkeley, CA, USA, 2005. USENIX Association.

[14] Garrett Hardin. The Tragedy of the Commons. *Science*, 162:1243–1248, 1968.

[15] Seung Jun and Mustaque Ahamad. Incentives in bittorrent induce free riding. In *P2PECON '05: Proceedings of the 2005 ACM SIGCOMM workshop on Economics of peer-to-peer systems*, pages 116–121, New York, NY, USA, 2005. ACM.

[16] John Nash. Non-cooperative games. *The Annals of Mathematics*, 54(2):286–295, 1951.

[17] Noam Nisan and Amir Ronen. Algorithmic mechanism design (extended abstract). In *STOC*, pages 129–140, 1999.

[18] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani. *Algorithmic Game Theory*. Cambridge University Press, New York, NY, USA, 2007.

[19] Larry Peterson, Tom Anderson, David Culler, and Timothy Roscoe. A Blueprint for Introducing Disruptive Technology into the Internet. In *Proceedings of HotNets–I*, Princeton, New Jersey, October 2002.

[20] Ryan Peterson and Emin Gün Sirer. Going beyond tit-for-tat: Designing peer-to-peer protocols for the common good. In *Proceedings of the Workshop on Future Directions in Distributed Computing*, 2007.

[21] Michael Piatek, Tomas Isdal, Thomas Anderson, Arvind Krishnamurthy, and Arun Venkataramani. Do incentives build robustness in bittorrent? In *NSDI'07*, Cambridge, MA, April 2007.

[22] Dongyu Qiu and R. Srikant. Modeling and performance analysis of bittorrent-like peer-to-peer networks. In *SIGCOMM '04: Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 367–378, New York, NY, USA, 2004. ACM.

[23] Jeffrey Shneidman and David C. Parkes. Rationality and self-interest in peer to peer networks. In *2nd Int. Workshop on Peer-to-Peer Systems (IPTPS'03)*, 2003.

[24] Ion Stoica, Robert Morris, David Liben-Nowell, David R. Karger, M. Frans Kaashoek, Frank Dabek, and Hari Balakrishnan. Chord: a scalable peer-to-peer lookup protocol for internet applications. *IEEE/ACM Trans. Netw.*, 11(1):17–32, 2003.

[25] Vivek Vishnumurthy, Sangeeth Chandrakumar, and Emin Gun Sirer. Karma: A secure economic framework for p2p resource sharing. In *Workshop on the Economics of Peer-to-Peer Systems*, Berkeley, California, 2003.

[26] Kevin Walsh and Emin Gün Sirer. Experience with an object reputation system for peer-to-peer filesharing. In *NSDI'06: Proceedings of the 3rd conference on 3rd Symposium on Networked Systems Design & Implementation*, pages 1–1, Berkeley, CA, USA, 2006. USENIX Association.

[27] Haifeng Yu, Michael Kaminsky, Phillip B. Gibbons, and Abraham Flaxman. Sybilguard: defending against sybil attacks via social networks. *SIGCOMM Comput. Commun. Rev.*, 36(4):267–278, 2006.