<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>A new analytical framework for studying protocol diversity in P2P networks</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Jin, X; Deng, J; Kwok, YK</td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td>The 2013 IEEE International Conference on Communications (ICC 2013), Budapest, Hungary, 9-13 June 2013. In IEEE International Conference on Communications, 2013, p. 2293-2297</td>
</tr>
<tr>
<td><strong>Issued Date</strong></td>
<td>2013</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10722/189862">http://hdl.handle.net/10722/189862</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>IEEE International Conference on Communications. Copyright © IEEE.; ©2013 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.; This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.</td>
</tr>
</tbody>
</table>
A New Analytical Framework for Studying Protocol Diversity in P2P Networks

Xin Jin\textsuperscript{1}, Jian Deng\textsuperscript{2}, and Yu-Kwong Kwok\textsuperscript{1}

\textsuperscript{1}Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam, Hong Kong
\textsuperscript{2}School of Microelectronics, Shanghai Jiao Tong University, Shanghai, China
Email: \{tojinxin,ykwok\}@eee.hku.hk, sheepjian@sjtu.edu.cn

Abstract—Thanks to years of research and development, current peer-to-peer (P2P) networks are anything but a homogeneous system from a protocol perspective. Specifically, even for the same P2P system (e.g., BitTorrent), a large number of protocol variants have been designed based on game theoretic considerations with the objective to gain performance advantages. We envision that such variants could be deployed by selfish participants and interact with the original prescribed protocol as well as among them. Consequently, a meta-strategic situation—judiciously selection of different protocol variants—will emerge.

In this work, we propose a general framework, Migration, based on evolutionary game theory to study the coevolution of peers for selfish protocol selection, and, most importantly, its impact on system performance. We apply Migration to P2P systems and draw on extensive simulations to characterize the dynamics of selfish protocol selection. The revealed evolution patterns shed light on both theoretical study and practical system design.

Keywords—Algorithms design, distributed systems, P2P networks, population coevolution, node rationality.

I. INTRODUCTION

The effectiveness of a peer-to-peer (P2P) network is contingent on the cooperation of autonomous participants. Such cooperation, or the lack of it, manifests as the result of aggregated actions taken by rational and strategic peers. Indeed, based on game theoretic considerations [1], [2], a plethora of protocol variants have emerged, aiming to take some unfair advantages in a P2P network. The most notable example is the BitTorrent file sharing networks. Apart from the original prescribed protocol [3], participating peers can choose to deploy protocol variants such as BitThief [4], BitTyrant [5], PropShare [6], BitMate [7], and Birds [2]. Consequently, a typical P2P network is anything but a homogeneous system from a protocol perspective.

In such a variegated environment, a number of interesting issues arise.

- Which protocol(s) is(are) more preferable in order to gain an edge in performance?
- Is there a protocol universally preferred by all peers?
- With the coexistence of diverse protocol variants, will a polymorphic equilibrium be achieved?
- If not, will the system become unstable or even chaotic?
- Most importantly, what is the impact of meta-strategic behaviors—dynamic protocol selection—on the entire system performance?

Unfortunately, existing research does not provide insightful answers to these challenging questions. Specifically, even with diverse design choices [1], [8], normal-form games are still being used for modeling pairwise encounters of peers. As such, general conclusions for systems with protocol diversity are largely unknown.

In this paper, we meet this research challenge by investigating strategic autonomous protocol selection in a variegated environment. Population coevolution is employed to describe the evolution of protocols taken by interacting strategic clients. In summary, our major contributions are two-fold:

- A Framework for Population Coevolution. We present a general framework, called Migration, in which the entire population in P2P networks is classified into different groups, each of which consists of peers taking the same protocol. Out of self-interests, peers dynamically migrate among diverse groups. In population coevolution, we study the dynamics and impact of such migration patterns.

- Validating Migration in P2P Systems. We validate Migration in P2P systems. Replicator dynamics [9], [10] are used for centralized population coevolution with the global knowledge of peer utility. We first formally demonstrate that both BitTorrent file sharing and P2P live streaming are population games, in stark contrast to traditional adoption of normal-form games. Extensive simulations on P2P live streaming show patterns of population coevolution.

The remainder of this paper begins with modeling popular P2P systems as population games in Section II. We introduce our framework for population coevolution in Section III. Section IV considers important implementation issues and presents extensive simulation results on P2P live streaming. Recent advances in incentives in distributed systems are stated in Section V. Section VI provides concluding remarks.

II. P2P SYSTEMS AS POPULATION GAMES

In this section, after a brief introduction to game modeling, we analytically prove that P2P systems for file sharing and live video streaming are both population games. Due to the limit of space, we omit analytical results associated with the models.
A. Game Modeling Preliminaries

Game theory is generally utilized to model and predict strategic interactions among distributed client peers [8], [9], [11].

1) Normal-form Games: Traditional studies often model peer interactions and node rationality as normal form games [1], [8], [11]. In normal-form games, pairwise encounters of peers determine the success of a strategy. That is, a peer determines its utility based on the other’s strategies. Denote by \( A = \{ a_i, 1 \leq i \leq K \} \) the strategy space (i.e. the set of all actions at peers’ choice). \( K \) is the number of protocols at nodes’ choice. \( u_{ij} \) is then the payoff for a pairwise encounter, in which one player takes the pure strategy \( a_i \), against another playing the pure strategy \( a_j \).

2) Population Games: In contrast, the success of a strategy in a population game depends on not the strategy of a neighboring counterpart but the strategy distribution of all concurrently online peers [9]. Without loss of generality, we study \( K \) types of peers, and \( m_i \) is the frequency of peers taking strategy \( a_i \), \( 0 \leq i \leq K \). Thus, a node that plays strategy \( a_i \) yields payoff \( f_i(m) \), where \( m \) denotes population composition, and \( \sum_{0 \leq i \leq K} m_i = 1 \).

B. BitTorrent-like file sharing as a population game

BitTorrent clients utilize optimistic unchocking to discover potentially better partners by offering some free-wins to some randomly selected peers. The original bandwidth allocation policy is to split one’s bandwidth equally among a few fastest uploading partners [3]. Strategic clients game to win by either exploiting the original optimistic unchocking design [4] or modifying the equal split policy [5], [6].

1) Game Against Optimistic Unchocking: We consider the coexistence of \( N_t \) BitThief nodes and \( N_r \) BitTorrent clients. The total number of clients is \( N = N_t + N_r \). The notorious BitThief [4] is customized to exploit free-wins in BitTorrent’s optimistic unchocking by increasing the number of local connections \( k_t \) with BitTorrent clients. Suppose that regular peers connect with both BitTorrent and BitThief clients. BitThief nodes exclusively connect with regular peers because their purpose is to exploit altruism in BitTorrent. Thus, for each optimistic unchock, BitThief can receive download rate \( C/(w_r + a_r) \) for 30 seconds [3], where \( w_r \) is the number of BitTorrent’s optimistic unchockes and \( a_r \) is the number of BitTorrent’s regular unchockes. Denote by \( k_r = a_r + w_r \) the total number of neighbors. Now we show that the number of optimistic unchockes \( X \) received by a BitThief client per round, which determines the download rate of BitThief clients, is dependent on the population composition.

**THEOREM 1.** The game between BitThief and BitTorrent is a population game because the expected number of optimistic unchockes received by a BitThief client is

\[
E[X] = w_r \cdot \min\left\{ \frac{k_t}{k_r} \left( 1 - \frac{a_r}{k_r} \right) \frac{N_r}{N_t} \right\},
\]

which depends on the population composition \( m = \{ N_r/N, N_t/N \} \).

**Proof.** Connection relationships between BitThief and BitTorrent clients dictate the following formula \( k_r \cdot \delta \cdot N_r = k_t \cdot N_t \), where \( \delta \) is the fraction of BitTorrent connections per BitThief client. From \( k_t \leq N_r \), we get \( \delta = \min\{1, \frac{N_t}{k_t} \frac{k_t}{N_r} \} \). To guarantee chunk dissemination among regular BT clients, the number of a BitTorrent client’s local connections with other BitTorrent peers is no smaller than the active set size \( a_r \). Thus, we have

\[
\delta = \min\{N_t, k_t \cdot N_r, 1 - a_r/k_r\}.
\]

Denote by \( n_u \) the number of opportunistic unchockes of a BitThief client by a BitTorrent client. Then,

\[
E[n_u] = w_r \cdot \delta,
\]

because \( \Pr\{n_u = i\} = C^i_{k_r} \left( \frac{w_r}{k_r} \right)^i \left( 1 - \frac{w_r}{k_r} \right)^{k_r-i} \).

From \( E[X] = \frac{w_r \cdot \delta \cdot N_r}{N_t} \) and \( k_t \leq N_r \), we get

\[
E[X] = w_r \cdot \min\left\{ \frac{k_t}{k_r} \left( 1 - \frac{a_r}{k_r} \right) \frac{N_r}{N_t} \right\}.
\]

2) Game Against Default Equal Split: Diverse types of bandwidth allocation schemes are proposed in the literature to game against equal split in regular BitTorrent [2], [5]–[7].

**THEOREM 2.** The coexistence of peers with different bandwidth allocation rules is a population game.

**Proof.** Consider two allocation rules \( g \) and \( f \) at peers’ choice. We study the steady state in which the average allocated bandwidth from peers with type \( x \) to active neighbors with type \( y \) is \( U_{xy} \). Then, for a type \( x \) peer, we have

\[
a_{xy} \cdot U_{xy} + a_{xf} \cdot U_{xf} = U,
\]

where \( a_{xy} \) is the number of active neighbors with type \( y \) for a type \( x \) peer, \( x, y \in \{ g, f \} \). Connections between peers with two allocation rules imply

\[
N_y \cdot a_{gy} = N_f \cdot a_{gf},
\]

where \( N_x \) is the number of peers with allocation rule \( x \in \{ g, f \} \).

The expected download rate of type \( x \) peer is thus

\[
D(U, x) = a_{xy} \cdot U_{xy} + a_{xf} \cdot U_{xf},
\]

where \( a_{gy} = a_gf \cdot \frac{N_y}{N_f} \), \( a_{gg} = k_r - a_{gf} \), and \( a_{ff} = k_r - a_{gf} \cdot \frac{N_f}{N_r} \). Given allocation rules, the optimal neighbor composition for a type \( x \) peer is determined by

\[
a_{xy}^* = \min\{\arg\max_{a_{xy}} D(U, x), \frac{N_y \cdot a_{xy}^*}{N_x}\},
\]

where \( x, y \in \{ g, f \}, x \neq y \).

Suppose that system converges to the optimal neighbor composition \( a_{xy}^* \). Take type \( f \) peer as example. From Equation 6, we get

\[
D(U, f) = a_{gf} \cdot \frac{N_g}{N_f} \cdot U_{gg} + (a - a_{gf} \cdot \frac{N_g}{N_f}) \cdot U_{ff},
\]

which depends on \( \frac{N_g}{N_f} \), and thereby the population composition of two types of peers. \( \square \)
C. Live Streaming Peer Selection as a Population Game

In P2P live streaming systems, overlay topology governed by peer selection can significantly affect streaming quality of peers [12]. In our study, clients are logically divided into multiple groups. In particular, peers with similar upload bandwidth, certain proximity in terms of round-trip time, or similar distances towards the source are clustered into the same group. We utilize clustering index to measure clustering effects. According to [13], the clustering index of peer $i$ is defined as

$$ c_i = \frac{k_i}{n_i}, $$

where $k_i$ is the number of neighbors within the residential group of peer $i$. The average clustering index for peers in group $k$ is

$$ C_k = \frac{\sum_{i \in N_k} c_i}{|N_k|}, $$

where $N_k$ is the set of clients in group $k$.

For clarity, we consider two extremes of coexisting peer selection strategies. One is traditional random peer selection, randomly selecting neighbors across the network; the other is network aware peer selection by choosing others in the same group for connection. Random strategy promotes a random graph, while network aware peer selection enforces clustering among peers within the same group. Different population composition of peers taking either peer selection will modify the extent of clustering. Our analytical results demonstrate the significant impact of clustering on streaming quality. This immediately validates our adoption of population games.

**THEOREM 3.** The coexistence of different peer selection strategies for P2P live streaming is a population game.

**Proof.** The chunk dissemination model is affected by [14], and we incorporate clustering index to investigate the interactions among node groups. In the following, we first show that clustering index is affected by population composition $m = \{m_r^i, m_n^i\}$, where $m_r^i$ is the fraction of random strategists and $m_n^i$ is the fraction of network aware strategists in group $i$. Then, we demonstrate that streaming quality is determined by $m$ and thereby prove that the coexistence of different peer selection strategies is a population game. We study the chunk distribution of a typical chunk $c$.

Denote by $N(t)$ the number of peers in the network at time slot $t$, and $N_i(t)$ the number of peers in group $i$ at time slot $t$. The clustering indexes of a random strategist and a network aware strategist from group $i$ are respectively $c_r = N_i(t)/N(t)$ and $c_n = 1$. Thus, the clustering index of group $i$ is

$$ C_i = m_r^i \cdot c_r + m_n^i \cdot c_n = m_r^i \cdot \frac{N_i(t)}{N(t)} + m_n^i, $$

determined by population composition $m$.

Next, given clustering index $C_i$ governed by population composition, we investigate the impact of clustering on streaming quality, evaluated by the number of chunk holders in group $i$ at time slot $t$, $X_i(t)$. The average distribution delay of peers in group $i$ is

$$ \overline{d_i} = \frac{\sum_{t=0}^{t_{max}} [X_i(t) - X_i(t-1)] \cdot t}{N(t_{max})}, $$

where $t_{max} = \arg\{X_i(t) = N_i(t)\}$. And the average distribution delay of the entire network is

$$ \overline{d} = \frac{\sum_{t=0}^{t_{max}} [X(t) - X(t-1)] \cdot t}{N(t_{max})}, $$

where $X(t) = \sum_{i \in G} X_i(t)$ is the total number of chunk holders in the network, and $t_{max} = \arg\{X_i(t) = N_i(t)\}$.

The average number of chunk holders for a peer in group $i$ is thus

$$ n_i(t) = \frac{1}{N_i(t)} \sum_{b=0}^{B} X_i(t). $$

To focus on topological dynamics, we assume random chunk scheduling. Denote by $Y_i(t)$ the number of chunk holders in group $i$ selecting $c$ to upload at time $t$. $k_i$ is the node degree of a peer from group $i$, and $k_i^l$ is the number of links with group $l$ for a peer from group $i$. $p_i(t)$ is the probability of chunk holding in group $k$. $A_i^l(t)$ is the number of connections from group $l$ yet without $c$ for a group $k$ node. $Z_{il}(t)$ represents the number of copies a peer from group $k$ can upload to group $l$.

$$ ∴ Y_i(t) = \frac{1}{n_i(t)} \cdot X_i(t) \quad \text{and} \quad p_i(t) = \frac{X_i(t)}{N(t)} $$

$$ ∴ A_i^l(t) \sim B(k_i^l, 1 - p_i(t)), $$

where

$$ k_i^l = \left\{ \begin{array}{ll}
    k_i \cdot C_i & \text{if } l = i \\
    (1 - C_i) \cdot k_i \cdot \frac{N_l(t) \cdot k_i (1 - C_i)}{\sum_{j \neq l} N_j(t) \cdot k_j (1 - C_j)} & \text{otherwise.} 
\end{array} \right. $$

$$ ∴ Z_{il}(t) = \min\{A_i^l(t), u_i \cdot \frac{k_i^l}{k_i}\}, $$

$$ ∴ X_i(t + 1) = X_i(t) + \sum_{l \in G} Y_i(t) \cdot Z_{il}(t), $$

which is affected by clustering index $C_k$ and with the initial condition:

$$ X_i(0) = u_s \cdot \frac{N_i(0)}{N(0)}. $$

Suppose that peers join and depart at the end of each time slot. System dynamics can be easily incorporated into the model:

$$ \{ \begin{array}{l}
    N(t + 1) = N(t) \cdot (1 - \mu_{\text{dep}}(t)) + \lambda_{\text{arr}}(t) \\
    X_i(t + 1) = \mu_{\text{dep}}(t) \cdot [X_i(t) + \sum_{l \in G} Y_i(t) \cdot Z_{il}(t)] \\
\end{array} \} \quad \text{for } i \in G, $$

where $\lambda_{\text{arr}}(t)$ is the aggregate joining rate of the system, and $\mu_{\text{dep}}(t)$ is the departure rate of each participating peer. □
III. Migration: A Framework for Population Coevolution

In this section, we propose Migration, a framework to study the coevolution of distributed nodes taking various strategies. It consists of a population game to model the interactions, and evolutionary dynamics for population coevolution. We have already introduced population games in Section II. Now, we explain population dynamics.

A. Population Dynamics for Global Coevolution

Evolutionary game theory is an extraordinary tool to investigate the behavioral dynamics in large populations. In essence, learning dynamics specify the transition rate from one group to another based on utility and population composition. Due to the lack of deterministic learning rules, converging to Nash equilibrium, we resort to stochastic learning. Due to the limit of space, we only introduce replicator dynamics [9].

**Replicator Dynamics.** The evolution of peer population follows the basic tenet of Darwinism:

\[ \frac{m_i}{m} = f_i(m) - \bar{f}(m). \]

where \( \bar{f}(m) = \sum_{0 \leq i \leq K} m_i \cdot f_i(m) \) is the average fitness. The rate of increase \( \frac{m_i}{m} \) of the group taking strategy \( a_i \) measures its evolutionary success. In this way, the replicator dynamics mimics natural selection. Such traditional evolutionary game theory investigates frequency dependent selection in well-mixed populations. At the same time, it is equivalent to fitness proportionate selection in the terminology of evolutionary computing [15].

IV. Evaluation

In this section, we evaluate our algorithms in P2P live streaming networks.

A. Implementation Issues

Network dynamics implies the volatile nature of utility experienced by distributed clients. Exponential moving average (EMA) is adopted to smooth fluctuations in a utility time series of node \( p \):

\[ \bar{U}_p(t) = \alpha \cdot u_p(t) + (1 - \alpha) \cdot \bar{U}_p(t-1), \]

where the constant \( \alpha \) reflects the importance given to the most recent data, and \( u_p(t) \) is the utility at time \( t \).

B. Evaluation for P2P Live Streaming

Network aware peer selection potentially boosts system performance for P2P live streaming [12]. That is, peers exclusively favoring higher upload capacity (capacity nerds), lower RTT (proximity nerds), smaller source hop count \( d_p \) (source distance nerds), and lower buffer map overlap (buffer map nerds), coexist and evolve according to replicator dynamics. The former three metrics can be easily defined. However, buffer map overlap cannot be directly derived before node connections. Thus, we utilize the difference between source hop counts of two nodes to discover peers with small buffer map overlaps because larger difference indicates higher probability that peers’ buffer maps are not strongly overlapped [16].
to this line of thought, Lin et al. [11] propose a cheat proof cooperation strategy by taking peer capacity into consideration. Measurements of PPLive users demonstrate the inefficacy of such bilateral incentive strategies, when applied to live video streaming environment [16]. To this end, Piatek et al. suggest to provide incentives by locating high contribution peers closer to the source servers. The rationale is that nodes closer to the streaming source can obtain higher streaming quality, as demonstrated by capacity aware overlay construction strategies. Feldman et al. [18] studies the evolution of cooperators, defectors, and reciprocators in P2P networks. However, all the above studies fail to consider decisions of individual peers for autonomous protocol selection. With the continual evolution of P2P systems, the advent of diverse protocol options shows the imperativeness of a study on the coevolution of distributed nodes in such a variegated environment.

VI. CONCLUDING REMARKS

Node rationality and incentives are both critical design considerations in distributed systems. However, the coevolution among strategic nodes for rational decisions is still largely unexplored. In this paper, we leverage on evolutionary game theory, present population coevolution algorithms, and reveal essential patterns. We first introduce population games and population dynamics in our framework Migration. In this manner, the elusive dynamical interactions among distributed nodes can be effectively scrutinized by investigating migration patterns of nodes taking different protocols. Extensive simulations for P2P live streaming are presented to validate our modeling efforts.

REFERENCES