Bitswap Analysis

David Grisham

In this document, we analyze 3 strategies for a simple 2-player Bitswap infinitely repeated game.

TODO: finish this intro. Might want to mention the paragraph below, but have to be clear that the first section is more general than the specific case done in the analysis

In a given round, a player's utility is dependent on the actions each of the players took in the previous round – in other words, the debt ratio only considers the immediately preceding round, rather than all previous rounds.

Actions and Utility Functions

A player has two possible actions: reciprocate (R) or defect (D). The utility function for player i at time t u_i^t:

$$u_i^t = \sum_{j \in \mathcal{N}_i} \delta_{a_j^t R} S_j(d_{ji}, \mathbf{d}_j^{-i}) B - \delta_{a_i^t R} B$$

where

- $\mathcal{N}_i$ is the neighborhood of user i (i.e. the set of peers i is connected to)
- $a_i^t \in \{R, D\}$ is the action user i takes in round t
- δij is the kronecker delta function
- d_{ji}^t is the reputation of user i as viewed by peer j (also referred to as the *debt ratio* from i to j) in round t
- $S_j(d_{ij}, \mathbf{d}_j^{-i}) \in \{0, 1\}$ is the *strategy function* of user j. This 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 bandwidth to give to peer i.
- $B > 0$ is the (constant) amount of bandwidth that a user has to offer in a given round. We make the simplifying assumption that the users are homogeneous in this value, so they all have the same amount of bandwidth to offer.

The terms *strategy* and *strategy function* are defined as:

- A *strategy* is meant in the standard game-theoretical sense, which is a predetermined set of actions that a user will take in a game (potentially dependent on that user's previous payoffs, the actions of its peers, etc.).
- A *strategy function* is a term used to specify the weighting function that a user uses when running the Bitswap protocol to determine how much bandwidth it wants to allocate to each of its peers whenever it's playing the R strategy.

Putting this all together, we see that the utility of peer i in round t is the total amount of bandwidth that i is allocated by its neighboring peers, minus the amount of bandwidth that i provides to its peers. If i reciprocates, then we say that they provide a total of B bandwidth to the network; otherwise (when i defects), i provides 0 bandwidth in that round.

We can write the debt ratio d_{ij} in terms of the number of bits exchanged between peers i and j:

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

where b_{ij}^{t-1} is the total number of bits sent from i to j from round 0 through round $t-1$ (so, all rounds prior to round t).

We can define b_{ij}^{t+1} in terms of b_{ij}^{t} and $\delta_{a_i^t R}$ as follows:

$$b_{ij}^{t+1} = b_{ij}^{t} + \delta_{a_i^t R} S_i(d_{ij}, \mathbf{d}_i^{-j}) B$$

So, the total number of bits sent from i to j increases by $S_i(d_{ij}, \mathbf{d}_i^{-j})B$ (the proportion of i's total bandwidth that i allocates to j) if and only if peer i reciprocated in round t (i.e., $a_i^t = R \implies \delta_{a_i^t R} = 1$).

Now we can write d_{ij}^{t+1} in terms of values from round t.

$$d_{ij}^{t+1} = \frac{b_{ij}^{t} + \delta_{a_i^t R} S_i(d_{ij}, \mathbf{d}_i^{-j}) B}{b_{ji}^{t} + \delta_{a_j^t R} S_j(d_{ji}, \mathbf{d}_j^{-i}) B + 1}$$

Analysis

The strategy function that user j will use to weight some peer i is:

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

$$u_i^t = \sum_{j \in \mathcal{N}_i} \frac{\delta_{a_j^t R} d_{ji}^t}{\sum_{k \in \mathcal{N}_j} d_{jk}^t} B - \delta_{a_i^t R} B$$

As this is an infinitely repeated game, we want to be able to calculate the discounted average payoff of player i, U_i.

$$U_i = (1 - \epsilon_i) \sum_{t=1}^{\infty} \epsilon_i^{t-1} u_i^t(\mathbf{a}^t)$$

where

- $\epsilon_i \in (0, 1)$ is the *discount factor*, which characterizes how much player i cares about their payoffs in future rounds relative to the current round.
- $\mathbf{a}^t$ is the vector containing each player's actions in round t. (**TODO**: define $\mathbf{a}^t$ with some sort of vector-comprehension notation, or just $(a_1^t, a_2^t, ...)$)

TODO: Mention that d_{ij}^t only considers previous round (and ensure explanation is consistent with analysis below)

We now consider 3 strategies: tit-for-tat, pavlov, and grim-trigger. For each of these, we'll determine whether the strategy is an subgame-perfect Nash equilibrium (SPNE) for the 2-player infinitely repeated game.

Tit-for-Tat

We start by analyzing the well-studed tit-for-tat (TFT) strategy. A peer that uses this strategy always takes the strategy that their peer took in the previous round. So, if player 1 plays action R (D) in round t, then player 2 will play action R in round $t + 1$ (D), and vice-versa.

To determine whether TFT is an SPNE for the 2-player infinitely repeated (simplified) Bitswap game, we will do the following:

- Consider an initial pair of actions at $t = 0$, (a_1^0, a_2^0).
- Assume player 1 plays TFT for all rounds.
- Consider two cases:
  1. Player 2 never deviates from TFT.
  2. Player 2 deviates from TFT for a single round, at $t = 1$, then plays TFT for all future rounds.
- Compare player 2's payoff for the infinitely-repeated game. If player 2's payoff in case 2 (U_2') is less than or equal to their payoff in case 1 (U_2) for all initial pairs of actions, then TFT is an SPNE. Mathematically, TFT is an SPNE if and only if $U_2' \leq U_2$ for all possible initial actions.

Let's look at the case where the initial pair of actions is (D, D).

(D, D) **at** $t = 0$

Here, both players start by playing the D strategy. We'll first consider the case where player 2 does not deviate from TFT. The strategies at each round then follow:

t	0	1	2	3	4	...
a_1^t	D	D	D	D	D	...
a_2^t	D	D	D	D	D	...

Since neither player deviates from TFT, they both continually play their opponent's previous strategy – the initial state is (D, D), so each player repeatedly plays D in this instance.

We can calculate the payoff of player 2 in this case – notice that, since neither player is ever giving or receiving, they payoff at each round is 0.

$$u_2^t = 0 \; \forall \; t \implies U_2 = 0$$

Now we consider this case where player 2 deviates from TFT for 1 round, at $t = 1$. The resulting action sequence is then:

t	0	1	2	3	4	...
a_1^t	D	D	R	D	R	...
a_2^t	D	R	D	R	D	...

In this case, player 2 deviates from TFT for a single round and plays R at $t = 1$ rather than D, then goes back to playing TFT for all rounds after that. Player 1, on the other hand, never deviates from TFT.

We can calculate player 2's discounted average payoff in this second case, U_2'. Note that we consider payoffs starting at $t = 1$, not $t = 0$ – this is because we're calculating the payoff that player 2 would perceive if they deviated from TFT for a single round, and that decision is happening at $t = 1$ in our case. (**TODO**: work on the explanation in this last sentence)

At $t = 1$, player 1 defects and player 2 reciprocates – since there are only two players (**TODO**: any other conditions, e.g. lookbehind length?), player 2 provides all of its bandwidth to player 1 when it plays R. Thus, player 1's payoff is $u_1^1 = B$ and player 2's payoff at $t = 1$ is $u_2^1 = -B$.

At $t = 2$, the players swap strategies (because they're back to playing TFT). So $u_1^2 = -B$ and $u_2^2 = B$.

The game alternates between these two states forever. Thus, the payoff for player 2 in this case is:

$$
\begin{aligned}
U_2' &= -B + \epsilon_2 B - \epsilon_2^2 B + \epsilon_2^3 B - \dots \\
&= B(-(1 + \epsilon_2^2 + \epsilon_2^4 + \dots) + \epsilon(1 + \epsilon_2^2 + \epsilon_2^4 + \dots)) \\
&= B(\epsilon - 1)\frac{1}{1 - \epsilon_2^2} \\
&= -\frac{B}{1 + \epsilon_2}
\end{aligned}
$$

Given U_2 and U_2', we can start to discern whether TFT might be an SPNE for this game. We see that $U_2' < U_2$ – this means that TFT *might* be an SPNE, but we have to verify that this is the case for all other initial conditions as well.

For the rest of the cases, we simply show the action sequences and the discounted average payoff results

(D, R)

When player 2 doesn't deviate from TFT:

t	0	1	2	3	4	...
a_1^t	D	R	D	R	D	...
a_2^t	R	D	R	D	R	...

Thus,

$$
\begin{aligned}
U_2 &= B - \epsilon_2 B + \epsilon_2^2 B - \dots \\
&= \frac{B}{1 + \epsilon_2}
\end{aligned}
$$

When player 2 does deviate from TFT:

t	0	1	2	3	4	...
a_1^t	D	R	R	R	R	...
a_2^t	R	R	R	R	R	...

Both players give and receive B bandwidth in all rounds $t \in [1, \infty)$, so:

$$
\begin{aligned}
U_2' &= 0 - \epsilon_2 0 + \epsilon_2^2 0 - \dots \\
&= 0
\end{aligned}
$$

Thus, $U_2' < U_2$ in this case.

(R, D)

When player 2 doesn't deviate from TFT:

t	0	1	2	3	4	...
a_1^t	R	D	R	D	R	...
a_2^t	D	R	D	R	D	...

Thus,

$$
\begin{aligned}
U_2 &= -B + \epsilon_2 B - \epsilon_2^2 B + \dots \\
&= -\frac{B}{1 + \epsilon_2}
\end{aligned}
$$

When player 2 does deviate from TFT:

t	0	1	2	3	4	...
a_1^t	R	D	D	D	D	...
a_2^t	D	D	D	D	D	...

$$
\begin{aligned}
U_2' &= 0 - \epsilon_2 0 + \epsilon_2^2 0 - \dots \\
&= 0
\end{aligned}
$$

We get a differently result in this case, namely $U_2' > U_2$. Therefore, **TFT is not an SPNE for this game**.

(R, R)

Even though we've already proven that TFT is not an SPNE, we present the analysis of the final case for completeness.

When player 2 doesn't deviate from TFT:

t	0	1	2	3	4	...
a_1^t	R	R	R	R	R	...
a_2^t	R	R	R	R	R	...

Thus,

$$U_2 = 0$$

When player 2 does deviate from TFT:

t	0	1	2	3	4	...
a_1^t	R	R	D	R	D	...
a_2^t	R	D	R	D	R	...

Thus,

$$U_2' = \frac{B}{1 + \epsilon_2}$$

We again get the result $U_2' > U_2$, which also indicates that TFT is not an SPNE.

Grim-Trigger

TODO

Pavlov

TODO

Discussion

TODO: compare results to that of prisoner's dilemma

TODO: discuss next steps (more peers)