Analysis and Attacks of decentralized content curation platforms

Andrés Monteoliva Mosteiro, Orfeas Stefanos Thyfronitis Litos, and Aggelos Kiayias

University of Edinburgh
a.monteoliva@serious.server, o.thyfronitis@ed.ac.uk, akiayias@inf.ed.ac.uk

Abstract. We will attack Steem.

1 Introduction

Steem is not incentive-compatible.

2 Related Work

Many people have done many similar things.

3 Model

1 Notation

- We denote the set of all probability distributions on set A as $\mathcal{D}(\mathcal{A})$.
- We denote the powerset of a set A with 2^A.
- $a \| b$ denotes the concatenation of a and b.

2 Properties of Post Voting Systems

A post voting system has the objective to arrange the posts according to the preferences of the participants. The ideal order is defined based on the likeability matrix for the posts.

Definition 1 (Post). *Let $N \in \mathbb{N}^*, L_i \in \mathcal{D}\left([0,1]^N\right)$. A post is defined as $p = (i, l(, v))$, with $i \in [N], l \sim L_i$ (and $v \in \mathbb{R}_+$).*

- ***Author.*** *The first element of a post is the index of its creator, i.*
- ***Likeability.*** *The likeability of a post is defined as $l \in [0,1]^N$ (, where l is drawn from L_i the Likeability Distribution of its creator u_i).*

– *(**Votes.** A post has an associated "vote" value, which is a real non-negative number. It is initialized at 0 and increases whenever a player votes for the post, as explained later in detail.)*

Let $P \in \mathbb{N}^*$ *the number of posts. Then* $\forall j \in [P]$, *let* $\text{creator}_j \in [N], l_j \sim L_{\text{creator}_j}$ *and* $p_j = (\text{creator}_j, l_j(,0))$. *The set of all posts is* $\mathcal{P} = \bigcup_{j=1}^{P} \{p_j\}$.[1,2]

Definition 2 (Post score). *Let post* $p = (m, l)$. *We define the* score *of* p *as* $\text{sc}(p) = \sum_{i=1}^{N} l_i$.

The score of a post is a single number that represents its overall worth to the community. By using simple summation, we assume that the opinions of all players have the same weight. In an ordered list of posts where higher posts are more visible, the "common interest" would require that a post with higher score appear before another post with a lower score.

Definition 3 (Ideal Post Order). *Let* $\mathcal{P}$ *a set of posts. We define* $\textsc{IdealOrder}(\mathcal{P})$ *as a list of the posts in* $\mathcal{P}$ *such that*

$$\forall i < j \in |\mathcal{P}|, \text{sc}(\textsc{IdealOrder}(\mathcal{P})[i]) \geq \text{sc}(\textsc{IdealOrder}(\mathcal{P})[j]) \quad .$$

[3]

Definition 4 (Player). *Let* $N \in \mathbb{N}^*$. *A player* u *is defined by her unique* $\text{id} \in [N]$, *her Likeability Distribution* $L \in \mathcal{D}\left([0,1]^N\right)$ *and her Strategy* $S \in \{H, G\} \times \mathbb{N}^* \times 2^{[N]}$. *Let the* i*th player be represented by the tuple* $u_i = (i, L_i, S_i)$. *The tuple of players is defined as* $\mathcal{U} = (u_1, \ldots, u_i, \ldots, u_N)$.

We will now explain each field in u_i *in detail:*

– ***Likeability Distribution.*** *The Likeability Distribution* $L_i \in \mathcal{D}\left([0,1]^N\right)$ *of* u_i *is a distribution on how likeable is the content produced by* u_i *to the rest of the players. The Likeability Distribution for the whole system is* $\mathcal{L} = (L_1, \ldots, L_i, \ldots, L_n)$.
– ***Strategy.*** *The strategy of* u_i *is defined as* $S_i \in \{H, G\} \times \mathbb{N}^* \times 2^{[N]}$, *where the first element is the player's core strategy, the second is her attention span and the third is her voting ring.*

[1] TODO: consider removing things in parentheses
[2] TODO: consider dropping L_i entirely
[3] TODO: discuss list notation

- ***Honest/Greedy.*** *H corresponds to the honest and G to the greedy strategy. An honest player u_i would vote a post p according to its likeability towards herself, $l_{p,i}$, that is to say she votes posts following her preferences. (For honest players, the value of the vote is computed as $v_{H,i} = a \cdot \mathcal{VP} \cdot l \cdot \mathcal{SP} + b$, where l is drawn from the Likeability distribution)*[4,5]*. In Steem terms, l can be understood as the weight of a vote.*
  A greedy player only votes for posts produced by users of its Voting Ring. (The value of vote for a player if u_i is greedy is defined as $v_{G,i} = a \cdot \mathcal{VP}_i \cdot \mathcal{SP}_i$,[6] *as in our model all greedy votes are executed with full weight.)*
- ***Attention Span.*** *This is a positive integer that represents the number of posts a player can consider voting simultaneously. For the benefit of simplicity, we will assume that this number is constant throughought all players.*[7]
- ***Voting Ring.*** *If player u_i is honest, her Voting Ring is $R_i = \emptyset$. If u_i is greedy, her Voting Ring is $R_i \in 2^{[N]}$. A voting ring is defined as $R_i = \{g_1, \ldots, g_j, \ldots, g_n\}$ where $g_j \in \mathcal{U}$ is the jth member of the voting ring and n is the size of the ring. Two greedy players will either have the same or disjoint voting rings ($\forall i \neq j \in [N], R_i = R_j \vee (R_i \cap R_j = \emptyset)$).*

The tuple of all strategies is defined as $\mathcal{S} = (S_1, \ldots, S_i, \ldots, s_N)$.

The set of players is defined as $\mathcal{U} = (u_1, \ldots, u_i, \ldots, u_N)$.

Definition 5 (Post-Voting System). *An algorithm F that accepts as inputs:*

- *list of posts (in some order)*
- *set of players*

All posts must have a likeability vector of length equal to the number of players.

The system returns an ordered list of the input posts, possibly reordered

Definition 6 (Honest Player v1). *A player u is considered* honest *if her utility is maximized when the result of the post-voting system equals her subjective ordering of the posts:*

$$\forall i < j \in |\mathcal{P}|, l_{u,u\text{-}\textsc{Order}(\mathcal{P})[i]} \geq l_{u,u\text{-}\textsc{Order}(\mathcal{P})[j]} \ .$$

[4] TODO: move to posts section
[5] TODO: explain a, b
[6] TODO: same
[7] TODO: discuss

[8]

Definition 7 (Honest Player v2). *A player u is considered* honest *if she always votes the best post of the top k (her attention span) with a weight equal to how much she likes the post.*[9]

Definition 8 (Convergence under honesty). *We say that a post-voting system F t-converges* under honesty *if, for every valid input $(\mathcal{P},\mathcal{U})$ where all players are honest, it is*

$$F\left(\mathcal{P},\mathcal{U}\right)_{1..t} = \textsc{IdealOrder}\left(\mathcal{P}\right)_{1..t} \quad .$$

Theorem 1. *The Steem system t-converges under honesty, assuming the following conditions on $N, P, \mathcal{SP}, v, att\_span, a, b, c$ ...*[10]

The above result is tight. If the conditions are violated the above theorem is not true.

3 Steem Execution

Algorithm 1 Each player creates one post

```
1: function GeneratePosts(𝒰)
2:     𝒫 = ∅                                   ▷ List of posts
3:     for u_i ∈ 𝒰 do
4:         l ←ʳ L_i                            ▷ Get likeability of posts
5:         𝒫 ← 𝒫 ‖ (i, l, 0)                   ▷ Add post to list of Posts
6:     end for
7:     𝒫 ← Shuffle(𝒫)                          ▷ Shuffle posts to a random order
8:     return 𝒫
9: end function
```

TODOS[11,12]

[8] TODO: discuss list notation
[9] TODO: discuss definition of honesty
[10] TODO: continue
[11] TODO: Improve argmax notation
[12] TODO: Write Greedy

Algorithm 2 Player votes for best of k posts

function VOTE($u_i, \mathcal{P}$)
  switch S_i **do**
    case *honest*
      $p_j \leftarrow \underset{p \in \mathcal{P}_{1..k}}{\operatorname{argmax}} \{l_{i,p}\}$
      Parse p_j as (m, l_p, v)
      $v' \leftarrow v + a \cdot \mathrm{VP}_i \cdot l_{i,p} \cdot \mathrm{SP}_i + b$
      $\mathrm{VP}_i \leftarrow \mathrm{VP}_i - (a \cdot \mathrm{VP}_i \cdot l_{i,p} + b)$
      $\mathcal{P} \leftarrow p_1 \| p_2 \| \ldots \| p_{j-1} \| (m, l_p, v') \| p_{j+1} \| \ldots \| p_N$
    end case
    case *Greedy*
      ▷ If post belongs to voting ring and not reached min VPower
      if $p \in s.R \wedge p.VPower > s.Min$ **then**
        $voteValue \leftarrow p.VPower \cdot weight \cdot sp$
        $p \leftarrow p.votes + voteValue$
      end if
    end case
  end switch
  return $\mathcal{P}$
end function

Algorithm 3 Players cast votes for r rounds

function CURATE($\mathcal{U}, \mathcal{P}, r$)
  for $j = 1$ to r **do** ▷ r voting rounds
    for $u_i \in \mathcal{U}$ **do**
      if ISVOTEROUND(j, S_i, r, N) **then**
        $\mathcal{P} \leftarrow$ VOTE $(u_i, \mathcal{P})$ ▷ Player i votes zero or one posts
      end if
    end for
    $\mathcal{P} \leftarrow$ ORDER $(\mathcal{P})$ ▷ Order posts by vote count after each round of votes
  end for
  return $\mathcal{P}$
end function

Algorithm 4 Calculates whether voting in this round is optimal

function ISVOTEROUND(j, S, r, N)

end function

Algorithm 5 Steem($\mathcal{U}, \mathcal{P}, r$)

1: $\mathcal{P} \leftarrow \textsc{Curate}\,(\mathcal{U}, \mathcal{P}, r)$
2: **return** $\mathcal{P}$

4 Results

Steem won't achieve high quality posts.

5 Further Work

Posts at any time

6 Conclusion

Keep inventing new decentralized content curation platforms.

7 Acknowledgements

We thank @seriousposter for their invaluable posts analyzing Steem and our mums for the cookies.

References