Digital Ecclesia: Towards an Online Direct-Democracy Framework

ABSTRACT
Citizens currently envision the transition to the online direct democracy. However, the existing frameworks support the participatory democracy. Inspired by the Athenians’ direct-democracy, we propose the initial version of the framework Digital-Ecclesia that offers dynamic citizens’ engagement for achieving their universal participation. Furthermore, the Digital Ecclesia makes fair decisions in the case of conflicting votes by applying a game-theoretic method. For achieving privacy and scalability in the decision-making, the architecture of the Digital Ecclesia is distributed, i.e., each node is modelled as an autonomous software-component. To evaluate the fairness of the decision-making in the Digital Ecclesia, we conduct experiments on a corpus of real-world citizens’ choices.

KEYWORDS
Distributed framework, message-passing model, conflicting choices, game theory.

ACM Reference format:

1 INTRODUCTION
Citizens increasingly seek electronic ways to cooperate for resolving their emerging issues, envisioning the transition from the traditional representative democracy to the online direct democracy. Citizens’ attempts have recently led to the adoption of software frameworks for online democracy by public organizations (e.g., open ministry in Finland, participatory budgeting in Paris) [1].

Limitations & challenges. A framework supports the direct democracy if it enables the universal and direct citizens’ participation [2], as was happening in Ecclesia (the sovereign governing body) of the Ancient Athenians’ democracy1 [3]. However, the existing frameworks support the participatory democracy, since they are used by citizens, who have previously asked their participation in specific working-groups. These frameworks do not ensure that relevant citizens are reached and prompted to engage (on-demand engagement). Towards supporting the direct democracy, one of the core challenges is the dynamic engagement of citizens.

However, the impact of the universal participation is that as the citizens’ number increases, the diversity of their choices usually grow, leading to conflicts. Ecclesia was guaranteeing the fair decision-making by compromising conflicting choices. On the contrary, the current frameworks make decisions assuming that single consensus (i.e., a single decision that represents the majority of citizens) can always be reached, violating the impossibility theorem for social decisions [4]. According to that theorem, single consensus on the common will cannot be manufactured out of citizens’ choices. In that case, we face the challenge of compromising conflicting choices making fair decisions, i.e., they represent the majority of citizens.

Contribution. To address these challenges, we propose the framework Digital Ecclesia, inspired by the Ancient Athenians’ democracy. The initial version of our framework offers dynamic citizens’ engagement towards achieving universal and direct participation. To do so, the Digital Ecclesia is modelled as a social network that offers high and dynamic reachability of citizens. In particular, citizens are dynamically engaged via receiving notifications for new working groups. Notifications are broadcasted in the network with encrypted messages, starting from nodes that offer high reachability of the network [5]. Since each network node notifies only its neighbours, our broadcast is time and message efficient.

In the case of high participation, the Digital Ecclesia makes decisions from a large number of citizens’ votes. For resolving privacy and scalability issues in the decision-making, the architecture of the Digital Ecclesia is distributed [6]. In particular, each node is an autonomous software-component that has local storage (e.g., votes are stored only locally) and executes in parallel the segments of the decision-making procedure that are related to locally-stored information. Thus, the Digital Ecclesia goes beyond the existing social-networking frameworks, whose nodes are simple user-profiles.

To make fair decisions, the Digital Ecclesia compromises conflicts by employing game-theoretic methods [7], which have not been adopted by the existing frameworks. According to [8], a decision compromises conflicting choices if it maximizes the Nash social-welfare function, assuming that citizens vote independently to each other and their choices are proportional to their preferences. To realize independent voting, we model the voting procedure as a non-cooperative game [9]. We further specify a distributed algorithm for employing the voting game. To ensure that citizens’ votes are proportional to their preferences, we finally propose a metric that assesses the semantic similarity of past voting topics.

Concluding, the current version of the Digital Ecclesia offers a distributed algorithm for supporting the two (communication & engagement and decision-making) out of the five key-functions need to be addressed in democracy frameworks [2]. To evaluate the fairness of the decision-making in the Digital Ecclesia, we conduct experiments on a corpus of real-world citizens’ choices.

The rest of the paper is structured as follows. Section 2 formally defines the notion of the voting game. Section 3 specifies our distributed framework. Section 4 presents the preliminary evaluation of our algorithm. Section 5 describes related approaches. Section 6 summarizes our contribution and discusses its future directions.

1It was the first time in the human history that the direct democracy was applied.
2 VOTING GAME

A game generally consists of a set of players (e.g., voters). Players make choices with a pay-off that depends on their utility functions [7]. We assume in the Digital Ecclesia voters provide a numerical choice that belongs to a scale of discrete values (e.g., 5-star scale). We further assume that the utility function of a voter is proportional to his/her (probabilistic) voting preferences. The latter is calculated based on voters’ past choices on similar voting topics. Overall, we define below the notion of a voting game \( G \) for a topic \( p \) as follows.

**Definition 2.1 (Voting game).** A voting game is defined by the tuple \( G = (T, N, R, U, W) \), such that:

- \( T = \{ t_1, t_2, \ldots \} \) is a finite set of textual terms for a topic \( p \);
- \( N = \{ 1, 2, \ldots \} \) is a finite set of voters;
- \( R = \{ 1, 2, \ldots \} \) is a finite set of numerical choices;
- \( U_{sr}(r) = k \times pr_{sr}(r) \) is the utility function of a voter \( n \) for a choice \( r \) and is proportional to his/her preferences \( pr_{sr} \) with respect to the choice \( r \) (\( k \) is the constant of proportionality);
- \( W(r) = \sum_{i=1}^{N} \log U_{sr}(r) \) is the value of the social-welfare function for a choice \( r \) over the whole set of voters. \( \square \)

**Definition 2.2 (Voting preference).** The preference \( pr_{sr} \) of a voter \( n \) on a choice \( x \) equals to the percentage of the times that s/he has made the choice \( x \) for a set \( s \) of similar topics over the total number of his/her all choices for these topics:

1. \( \forall r \in R, pr_{sr}(r) = \frac{\sum_{i=1}^{s} sim(t_i, f(t_r))}{\sum_{i=1}^{s} sim(t_i, f(t_k))} \in [0, 1] \),
   - \( v_{sr}(s) \) is the total number of times that a voter \( n \) has made a choice \( x \) for a set \( s \) of similar topics;
   - \( c \), \( s = \{ p_1, p_2, \ldots, p_{|s|} \} : \forall p_i \in s, sim(p_i, p_1) > \theta \ | \theta \in [0, 1] \), is the set of the past topics \( p_i \) whose similarities to the current topic \( p \) are greater than a threshold \( \theta \);
   - \( d \), \( sim(p_i, p_j) = \frac{\sum_{i=1}^{k} sim(t_i, f(t_k))}{|T_1|} \), \( t_k \in T_1 \land f(t_k) \in T_2 \land |T_1| \leq |T_2| \) is the semantic similarity among the terms \( T_1 \) and \( T_2 \) of two topics \( p_i \) and \( p_j \):
     - i. \( f : t_k \rightarrow t_m \ | \sum_{k=1}^{n} sim(t_k, f(t_k)) \) is max;
     - ii. \( sim(t_k, t_m) = sim_{lin}(t_k, t_m) \).
2. \( \sum_{r=1}^{R_{max}} pr_{sr}(r) = 1 \), the sum of the rating probabilities of a consumer in the rating space \( R \) equals to 1. \( \square \)

Concerning the metric \( sim(p_i, p_j) \), it calculates the similarity between two topics as the average of the similarities of their most similar pairs of terms. To do so, the metric solves the assignment problem \( f \) of the maximum-weighted matching in a bipartite graph [10], where the nodes correspond to terms and the edges to their similarities. Regarding the metric \( sim_{lin}(t_k, t_m) \), it assesses the relatedness of terms by calculating the Lin’s similarity of the corresponding ontology concepts in WordNet [11].

3 DISTRIBUTED FRAMEWORK

As previously discussed, the Digital Ecclesia is modelled as a bidirectional network of software components that communicate to each other by passing encrypted messages. Each component corresponds to a single citizen, who holds authentication credentials. The topology of the network (a.k.a. communication links) can be dynamically updated during its lifetime. We call neighbours the nodes connected with a direct communication-link.

To offer dynamic engagement and distributed voting game, the Digital Ecclesia includes the distributed algorithm in Section 3.1, whose time and message complexity are discussed in Section 3.2.

### 3.1 Distributed Algorithm

Our algorithm (Alg. 1) is triggered by a node (hereafter called source) that corresponds to one of the citizens who initiated a new working group. The algorithm continues to the voting procedure only if the source offers high reachability of the network. To cope with the dynamicity of the network (e.g. communication links may have been updated), the reachability of the source is dynamically calculated, as explained in the following.

The algorithm consists of a *local program* for each node that provides the ability to nodes to perform local computations and send/receive messages. Since nodes do not usually have the same computation clock, messages are sent *asynchronously*. Nodes may send/receive the messages, \(< \text{reach} >, < \text{parent} >, < \text{source} >, < \text{vote} >, \) and \(< \text{utility} > \) (Alg. 1 (1, 9, 11, 17, and 32)).

**Reach message.** A reach message is initially sent by the source node to all its neighbours (Alg. 1 (2-4)). Every node that receives the message sets its parent node, sends parent and source messages (see below), and forwards the reach message to its neighbours except its parent (Alg. 1 (6-8)). Thus, a reach message is gradually forwarded (neighbour-by-neighbour) to all the nodes that are reachable from the source node (*broadcast* in message-passing systems [12]). To broadcast a reach message in a controlled way (i.e. messages do not wander in the network via different paths), a node does not forward messages that have already received (Alg. 1 (5)).

**Source and parent messages.** Since every node maintains its parent and children (distributed fashion), the algorithm constructs a *minimum spanning-tree* [13] (rooted at the source node) of the network part that is reachable from the source node. Moreover, every source-message is gradually forwarded back to the source node (Alg. 1 (16) – *convergecast* in message-passing systems [12]). By receipt of source message, the source node increases the number of the received messages (Alg. 1 (13)). If the number is greater than a threshold \( \omega \), then the source node has high reachability, and sends vote messages to its children (Alg. 1 (14-15)). Note that the source node does not send vote messages until the time limit has been expired in order to avoid to send them multiple times.

**Vote message.** When a node receives a vote message, then the node forwards the message to its children (Alg. 1 (18)). Following, the node asks by the citizen to vote for the current topic (Alg. 1 (21)), provided his/her credentials (Alg. 1 (19)). If the authentication is not successful, then the execution of the node for this topic is stopped (Alg. 1 (20)). Otherwise, the current vote of the citizen is locally stored, meeting our *privacy* requirement (Alg. 1 (29)). Following, the node updates the voting preferences of the citizen (Alg. 1 (23-28)) – applying Def. 2.2 – and sends the set of its

---

1. We leave as future work the topology management of the Digital Ecclesia.
2. For practical reasons, we considered in the implementation of the Digital Ecclesia upper bounds on voting time and message delays.
3. In the implementation of the Digital Ecclesia, we used the threshold value 0.6.
Algorithm 1 Dynamic Engagement & Distributed Voting Game

\[\text{Code for node } n_i, 1 \leq i \leq M \text{ (the total number of network nodes).}\]

\begin{algorithm}
\caption{Dynamic Engagement & Distributed Voting Game}
\begin{algorithmic}
\State \textbf{Input: } \(n_i, \theta, \omega\)
\State 1. \textbf{upon receiving} < reach, p, R > from neighbour \(n_j\) \textbf{do}
\State 2. \quad \textbf{if } \| R \| = \| n_j \| \textbf{ then} \quad \text{It corresponds to the source node.}
\State 3. \quad \quad \text{parent } \leftarrow n_j
\State 4. \quad \quad \text{SEND} < \text{reach}, p, R > \textbf{ to all neighbours}
\State 5. \quad \textbf{else if } \text{parent} = \text{null} \textbf{ then}
\State 6. \quad \quad \text{parent } \leftarrow n_j
\State 7. \quad \quad \text{SEND} < \text{parent} > \textbf{ and} < \text{source} > \textbf{ to parent}
\State 8. \quad \quad \text{SEND} < \text{reach}, p, R > \textbf{ to all neighbours except } n_j
\State 9. \quad \textbf{upon receiving} < \text{parent} > \textbf{ from neighbour } n_j \textbf{ do}
\State 10. \quad \quad \text{children } \leftarrow \text{children} \cup \{n_j\}
\State 11. \quad \textbf{upon receiving} < \text{source} > \textbf{ do}
\State 12. \quad \quad \text{++reached}
\State 13. \quad \quad \textbf{if } \text{reached} > \omega \text{ and } \text{expired} = \text{true} \textbf{ then}
\State 14. \quad \quad \quad \text{SEND} < \text{vote}(p, R, \theta) > \textbf{ to children}
\State 15. \quad \textbf{else } \text{SEND} < \text{source} > \textbf{ to parent}
\State 16. \quad \textbf{upon receiving} < \text{vote}(p, R, \theta) > \textbf{ do}
\State 17. \quad \quad \text{SEND} < \text{vote}(p, R, \theta) > \textbf{ to children}
\State 18. \quad \quad \text{correct } \leftarrow \text{AUTHENTICATE}
\State 19. \quad \quad \textbf{if } \text{correct} = \text{false} \textbf{ then} \text{RETURN } \text{end if}
\State 20. \quad \quad r \leftarrow \text{VOTE}(p, R, \theta)
\State 21. \quad \quad \text{STORE}(p, r)
\State 22. \quad \quad \text{for all } p_i \in (R_n - \{p\}) \textbf{ do}
\State 23. \quad \quad \quad \text{if } \text{SIM}(p, p_i) > \theta \textbf{ then}
\State 24. \quad \quad \quad \quad s \leftarrow s \cup \{p_i\}
\State 25. \quad \quad \quad \quad \quad \text{++}u_r(s)
\State 26. \quad \quad \quad \quad \quad \text{if } r = R_n(p_i) \textbf{ then} \quad \quad \quad \text{++}u_n(r, s)
\State 27. \quad \quad \quad \quad \quad \quad \quad \quad \text{pr}_n(r) \leftarrow \text{pr}_n(r) + \text{pr}_n(s)
\State 28. \quad \quad \quad \quad \quad \text{STORE}(\text{pr}_n(r))
\State 29. \quad \quad \text{\langle utility(\text{pr}_n) \rangle } \leftarrow \text{ENCRYPT(\text{pr}_n)}
\State 30. \quad \quad \text{SEND } \langle \text{utility(\text{pr}_n)} \rangle \textbf{ to parent}
\State 31. \quad \textbf{upon receiving} < \text{utility(\text{pr}_n)} \rangle \textbf{ from node } n \textbf{ do}
\State 32. \quad \quad \text{if } n_i = n \textbf{ then}
\State 33. \quad \quad \quad \text{pr}_n \leftarrow \text{DECRYPT(< utility(\text{pr}_n) >)}
\State 34. \quad \quad \quad \text{utilities.add(pr}_n\text{)}
\State 35. \quad \quad \quad \textbf{if } \text{expired} = \text{true} \text{ or } \text{\text{\mid utilities\text{}}} \leq \text{reached} \textbf{ then}
\State 36. \quad \quad \quad \quad \text{for all } r \in R \textbf{ do}
\State 37. \quad \quad \quad \quad \quad W(r) \leftarrow \sum_{u_i} |u_i| \log U_i(r)
\State 38. \quad \quad \quad \quad \quad \text{if } W(r) > \max \textbf{ then}
\State 39. \quad \quad \quad \quad \quad \quad \max \leftarrow W(r)
\State 40. \quad \quad \quad \quad \quad \quad \text{r}\_\text{max} \leftarrow r
\State 41. \quad \textbf{else } \text{SEND } < \text{utility(\text{pr}_n)} \rangle \textbf{ to parent}
\end{algorithmic}
\end{algorithm}

Utility message. A node that has received a utility message forwards the message to its parent. All the utility-values are gradually collected to the source node (convergecast). The source node decrypts each received utility-message (Alg. 1 (34)) and maintains the utility values (Alg. 1 (35)). Before making a decision, the source node waits until a time limit has expired or all the reachable nodes have sent utility messages (Alg. 1 (36)). Finally, the source node makes a decision that maximizes the social-welfare function (Alg. 1 (37–41)) – applying Def. 2.1.

### 3.2 Time and Message Complexity & Scalability

In the broadcast of reach messages, the time complexity is \(O(2 \times N)\), since all the nodes receive a reach message at most twice. The message complexity is \(N\), since every node sends its own reach message. In the convergecast of source messages, since all the nodes send their source messages to their parents based on a known spinning-tree, the time complexity is \(\log N\), where \(\log N\) is the depth of the tree. The message complexity is \(N\), since every node sends its own source message. In the broadcast of vote messages, the time and message complexity are \(N\), since all the nodes receive a vote message exactly once. In the convergecast of utility messages, the time complexity is \(\log N\) and the message complexity is \(N\). The overall (time and message) complexity of the algorithm, \(O(\log N + 2\log N) = O(N)\), asymptotically scales in a linear way to the total number \(N\) of the nodes that are reachable from the source node.

### 4 PRELIMINARY EVALUATION

**Dataset.** To evaluate our algorithm on large-scale real-world votes, we used the publicly available Amazon numerical ratings on thousands of products [15]. We considered that each product corresponds to a voting topic and a rating corresponds to a vote. The products had been organized into high-level categories. Each category contains products that may be similar to each other in terms of their descriptive terms. Due to lack of space, we indicatively present the results for five middle-sized categories (Table 1).

**Evaluation methodology.** We implemented a graph simulation of a distributed network and executed our algorithm for each product of the reported categories. At each execution, we configured the network to contain all the consumers who rated an examined product. We also considered a topology of the network such that all the nodes are reachable from at least one source node. To evaluate the fairness of decisions, we present the results for products that have conflicting ratings, i.e. there is no rating for a product that represents the majority of consumers. We compared the fairness of the decisions made by our algorithm against the arithmetic mean and median of ratings. These aggregators are the most widely used for making decisions on consumers’ choices [16]. We assume that a decision is fair if it represents at least the 60% of the consumers’ preferences (i.e. minimum percentage for majority decisions).

### Table 1: Statistics for the reported categories of products.

<table>
<thead>
<tr>
<th>ID</th>
<th>Category</th>
<th>#consumers</th>
<th>#products</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>Beauty</td>
<td>1210271</td>
<td>249274</td>
</tr>
<tr>
<td>A2</td>
<td>Tools &amp; Home Improvement</td>
<td>1212468</td>
<td>260659</td>
</tr>
<tr>
<td>A3</td>
<td>Video Games</td>
<td>826767</td>
<td>50210</td>
</tr>
<tr>
<td>A4</td>
<td>Toys &amp; Games</td>
<td>1342911</td>
<td>327698</td>
</tr>
<tr>
<td>A5</td>
<td>Health &amp; Personal Care</td>
<td>1851132</td>
<td>252331</td>
</tr>
</tbody>
</table>

---

6. [jmcauley.ucsd.edu/data/amazon](https://jmcauley.ucsd.edu/data/amazon)
7. We have reached to analogous conclusions for the remaining categories.
8. [https://www.princeton.edu/](https://www.princeton.edu/)
We observe that the percentages of the fair decisions made by our work Digital-Ecclesia that offers dynamic citizens’ engagement and works (e.g. Open Town Hall, Agoravoting, etc.) range from the 76% to the 81% of products. On the contrary, the corresponding percentages for the arithmetic mean (resp. median) range from the 33% (resp. 30%) to the 46% (resp. 40%) of products. These percentages further show that the decisions made by our algorithm are not the same with arithmetic means and medians in at least the 30% of products (esp. from the 30% to the 45% of products). Concluding, our algorithm makes fair decisions in a high number of cases and is more effective than the typical aggregators.

5 RELATED WORK

From architecture perspective, we organize the existing frameworks for online participatory democracy in three generations: (i) centralized Web-based frameworks (e.g. [17]), (ii) cloud-based frameworks (e.g. Open Town Hall\(^{\text{b}}\)), and service-oriented frameworks (e.g. [2, 18, 19]). Among the five key-functions [2], the existing frameworks have primarily focused on proposal making & voting (e.g. Agora Voting\(^{\text{a}}\)), online collaboration & discussion (e.g. Etherpad\(^{\text{c}}\)), and social communication [19].

Towards supporting the direct democracy, we faced the challenges of the dynamic engagement of citizens and distributed voting. Regarding these challenges, the existing frameworks model citizens by simple user-profiles, without ensuring that relevant citizens are reached (on-demand engagement). Moreover, citizens’ data are stored on the server side, raising security and privacy issues. We further faced the challenge of conflicts that emerge in a large number of citizens’ votes. Concerning this challenge, the existing frameworks make single-consensus decisions, assuming there are decisions that can represent the majority of votes. Single-consensus decision-making models that have been mainly used are decision trees [20] and Bayesian networks [21].

6 CONCLUSIONS AND FUTURE WORK

To support the direct democracy, we proposed the distributed framework Digital-Ecclesia that offers dynamic citizens’ engagement and distributed game-theoretic decisions. We evaluated our framework on a corpus of real-world choices and the preliminary results showed that our algorithm makes fair decisions in a high number of cases.

A future research-direction is to extend our distributed algorithm with sophisticated security and privacy techniques. On top of this, the voting procedure could be modelled as a cooperative game that gives incentives and rewards to citizens in order to make better choices. A final direction is to extend our framework to support the remaining key-functions of online democracy-frameworks.

REFERENCES


\(^{\text{a}}\)http://opentownhall.com
\(^{\text{b}}\)http://agoravoting.org
\(^{\text{c}}\)http://etherpad.org

![Figure 1: Products with conflicting ratings and fair decisions.](image-url)

Evaluation results. Fig. 1 depicts for each category the percentages of the products with both conflicting ratings and fair decisions. We observe that the percentages of the fair decisions made by our algorithm range from the 76% to the 81% of products. On the contrary, the corresponding percentages for the arithmetic mean (resp. median) range from the 33% (resp. 30%) to the 46% (resp. 40%) of products. These percentages further show that the decisions made by our algorithm are not the same with arithmetic means and medians in at least the 30% of products (esp. from the 30% to the 45% of products). Concluding, our algorithm makes fair decisions in a high number of cases and is more effective than the typical aggregators.