Improving On-Line Debates by Aggregating Citizen Support[*]

Maite LOPEZ-SANCHEZ [a,1], Marc SERRAMIA [b,a], and
Juan A. RODRÍGUEZ-AGUILAR [b]

[a] *Universitat de Barcelona (UB)*
[b] *Institut d'Investigació en Intel·ligència Artificial (IIIA-CSIC)*

Abstract. Currently, Digital Democracy is gaining momentum thanks to online participation platforms, which have emerged as innovative tools that enable citizens to participate in decision making processes. Through these tools, participants can issue proposals and engage into debates by both stating arguments in favour or against and/or by supporting other people's arguments. In this paper we propose a new support aggregation method derived from the combination of two complementary aggregation methods previously introduced. Additionally, we propose a resilience metric for measuring the quality of the aggregated opinion. We apply our contributions to debates conducted in the Decidim participatory platform.

Keywords. AI applied to participation, information fusion, debate quality.

1. Introduction

Most currently established democracies follow a *representative democracy* model where people periodically elect persons to represent them. However, it is often the case that people feel disconnected from their representatives, who retain the power during the inter-election periods. Against this, participatory portals are designed to bridge the gap between citizens and their governments. They foster participation in the public sphere by supporting the co-creation of proposals, public deliberation or joint decision-making. This supposes a transformation of the traditional model of representative democracy into a model of participatory democracy [19,7] where citizens become involved in decisions about some of the issues that most affect them, that is, the public interest [17]. This transition, in addition to mitigating the disaffection that a part of the citizenry feels towards politics, can lead to a substantial improvement in the acceptance of the decisions taken.

Participatory portals can also help to face challenges and crises (e.g., sustainable development goals, climate change, or pandemics) in a more cohesive way by enabling informed and reasoned decision making processes and by helping to mitigate the increasing polarisation of societies. In fact, debate has been proposed as an antidote to polarisation

Figure 1. Different initiatives to face COVID-19 pandemic enabled by the Decidim platform.

from some social actors such as the initiative of Spaceship Media [10], which emerged in the United States in the context of the political fracture produced in the Trump era. They argue that dialogue from difference is essential for the proper functioning of democracy.

We find numerous platforms for citizen participation. Among these, it may be worth mentioning: the French initiative Purpoz (previous "Parlement et Citoyens"[8]), where we can nowadays find debates on the climate bill or vaccination; the Petitions UK platform [13] in the United Kingdom, where 441 petitions got a response from the Government and 71 petitions were actually debated in the House of Commons by May 2021; or Rousseau in Italy [15] named after the theorist *par excellence* of participation [12]. Interestingly, some of the participatory democracy platforms –such as Decidim [6] or Consul [5]– emerged from popular movements, are open source, and are being used by local governments (e.g. the Generalitat de Catalunya [11]) and numerous municipalities such as Madrid, Reykjavik [1], New York, Buenos Aires or Barcelona [4]. Indeed, their use is not limited to public administrations, since organizations such as universities (e.g., Universidade da Coruña) or cooperatives (e.g., Som energia) are also using them to make their governance model more participatory.

In fact, although the pandemic has further promoted the digital divide among the most disadvantaged population, it has also turned out to be an indisputable catalyst for the digitalization of most of the population. In this manner, it poses a unique opportunity for participatory tools to demonstrate that the required dedication effort is worthy because people can feel part of the community and its decision process. For instance, Figure 1 illustrates that Decidim [6] was readily used in different initiatives facing the COVID-19 pandemic. If this opportunity is seized, the gain in transparency, traceability and accountability, as well as inclusion (both social and gender inclusion, since we need to transfer the role of women that has traditionally been relegated to the domestic sphere, to open it to participation in public spaces) represents a qualitative leap forward in our current democratic models.

Deliberation is a key mechanism to guarantee the quality of the decisions made. Indeed, debates in participatory platforms can facilitate informed and reasoned decision making when including the advantages and disadvantages of each proposal. However, the further the discussion progresses (or the larger the number of participants), the more complex to follow the argumentation. That is why it is vitally important to structure

these comments in such a way that the arguments that are for and against the proposal are clearly distinguished. In doing so, if participants then express their opinion about these arguments, then information aggregation operators can be applied to combine the different opinions into a single value that represents a joint support to the proposal.

This paper builds upon two aggregation methods –PAM [14] and TODF [9]– that compute this support (or joint opinion) in a more sophisticated way than the typical majorities, which suffer from a certain tendency to centralization. On the one hand, PAM [14] can consider aspects such as the number and importance of the expressed opinions to discard weak arguments that may hinder the resulting aggregated opinion, or even identify cases in which there is insufficient information to correctly evaluate a proposal. On the other hand, TODF [9] exploits the debate structure and guarantees properties –such as anonymity or non-authoritarianism– that have typically been considered within Social Choice Theory. We propose a hybrid aggregation method that inherits the advantages of both PAM and TODF methods and is tailored for the Decidim platform. We analyse these alternative methods by comparing the resulting aggregated support for some instances of real debates from Decidim Barcelona [4].

Furthermore, we propose a resilience measure to assess the quality of the decision taken w.r.t the proposal (i.e., if it is accepted or rejected). We define the resilience by considering the modifications that would be necessary to reverse the decision. Hereby, the more changes needed, the more resilient the decision.

We structure the paper as follows. Next section briefly introduces PAM and TODF aggregation methods so to provide the basis for our hybrid method, which is discussed in Section 3. Subsequent Section 4 proposes the resilience measure and illustrates its usage for some proposals from Decidim Barcelona. Finally, Section 5 concludes the paper and discusses possible future paths for research.

2. Background

This section is devoted to briefly introduce the two opinion aggregation methods: the *Proposal Argument Map* (PAM) and the *Target oriented discussion framework (TODF)*.

For illustration purposes, Figure 2 a) shows proposal[2] 50 from Decidim Barcelona [4]. It received 57 direct users' supports and aroused a debate consisting of 22 (in favor, against and neutral) arguments. Figure 2 b) illustrates how PAM groups arguments in favour and against, indicates how many citizens expressed their opinions over each argument, and represents the aggregated opinions using a 5-star scale. Figure 2 c) depicts the overall argument structure in the debate of proposal 50 (τ) as represented by TODF. Circles linked with green arrows represent arguments in favour, whereas arguments against are linked with red arrows. TODF signals accepted, undecidible, and rejected arguments (as computed by it) with green, yellow, and red colours respectively.

2.1. Proposal Argument Map (PAM)

Briefly, a proposal argument map (PAM) is a structure representing the content of a debate in terms of those arguments in favour and those arguments against a given proposal. Formally, Rodriguez-Aguilar et al. [14] define it as:

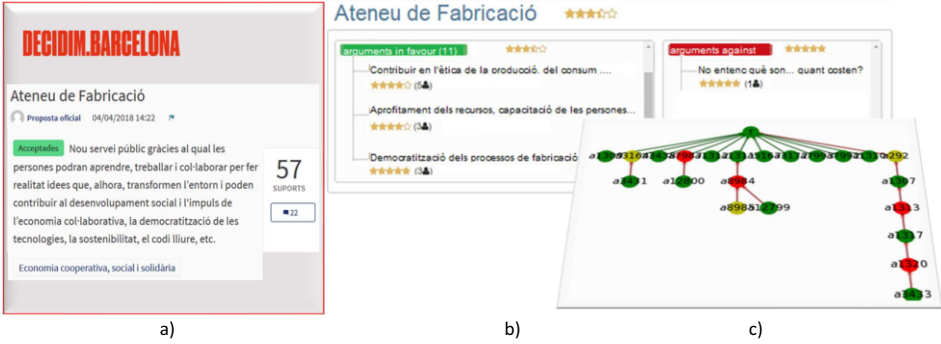[2]Proposal 50 (Pla d'Actuació Municipal 2016-2019): https://decidim.barcelona/pam/proposals/ateneu-de-fabricacio, https://github.com/elaragon/metadecidim/blob/master/proposals/00050.json

Figure 2. Proposal 50 as in a) Decidim (just general description shown), b) PAM, c) TODF.

Definition 1. *A Proposal Argument Map (PAM) is a triple $\langle p, A_p, \kappa \rangle$, composed of a proposal p, an argument set A_p and a function κ that classifies the arguments (as being in favour, neutral, or against the proposal).*

κ groups arguments in two separate sets: $A_p^+ = \{a \in A_p | \kappa(a) = 1\}$ includes those arguments in favour of the proposal and $A_p^- = \{a \in A_p | \kappa(a) = -1\}$ groups those against. Figure 2 b) depicts A_p^+ on the left column and A_p^- on the right. Neutral arguments (those with $\kappa(a) = 0$) are disregarded since they do not contribute much to the final decision.

An argument $a \in A_p$ is a pair $a = (s, O_a)$, where s is the argument description and O_a is the set of issued opinions. We then can combine these opinions into a single value rating the argument. We refer to this opinion aggregation as argument support ($S_{arg}(a)$), and compute it as a weighted mean. Weights in this function are computed by considering the opinions expressed using a 5-star scale, their number (see Figure 2 b)) and an importance opinion function meant to favour strong (clear) opinions over neutral positions —which can be associated with indecision (see [14] for further details).

This opinion support allows us to discard weak (non-relevant) arguments that would hinder the aggregated support. We consider an argument to be *relevant* if it has a significant number of opinions and it has enough support (significance is computed in terms of proportion of the issued opinions w.r.t. the maximum number of opinions issued for the other arguments in the debate).

Hence, PAM just focuses on relevant arguments in both A_p^+ and A_p^- and scales up the computation of the opinion support for the resulting sets $A_p^{R^+}$ and $A_p^{R^-}$. In this case, the aggregation operator we use is the WOWA (Weighted Ordered Weighted Average) operator [18] which, in addition to the importance of aggregated opinions over (relevant) arguments, it also considers their relative ordering (again, see [14] for a detailed explanation). We denote the support of these two argument sets as $S_{set}(A_p^{R^+})$ and $S_{set}(A_p^{R^-})$ respectively. However, it is worth noticing that we only compute the aggregated support of an argument set whenever there are relevant arguments. Otherwise, we consider we lack enough quality opinions to compute the aggregated opinion.

Finally, given a proposal p and the supports of the (non-empty) sets of relevant arguments in favour and against the proposal (i.e., $S_{set}(A_p^{R^+})$ and $S_{set}(A_p^{R^-})$ respectively), we compute the proposal support $S_{prop}(p)$ by applying the same WOWA operator (details can be found in [14]).

2.2. Target oriented discussion framework (TODF)

The Target Oriented Discussion Framework (TODF) [9] is an alternative support aggregation method to be used in debates. TODF focuses on the debate structure, which is specified in terms of arguments that can attack or defend both the proposal as well as other arguments.

Formally, a Target Oriented Discussion Framework is a structure $TODF = \langle \mathscr{A}, \mapsto, \vdash, \tau \rangle$, where $\mathscr{A}$ is a set of arguments; $\mapsto \subseteq \mathscr{A} \times \mathscr{A}$ is an attack relation (if $a \mapsto a'$, then a is attacking a'); $\vdash \subseteq \mathscr{A} \times \mathscr{A}$ is a defence relation (if $a \vdash a'$, then a is defending a') and τ is the target (in our case the proposal). Figure 2 c) shows how we represent attack and defence relations with red and green arrows respectively.

In our Target Oriented Discussion Framework, citizens express their opinions by assigning qualitative values (labels) to arguments and the proposal (which is a particular argument). Such labels are: in , meaning they agree with the argument; out, if they disagree; and undec, if they are neutral, not sure, or not defined. Each citizen i provides an argument labelling L_i, so we can define $\text{in}_L(\mathscr{A}) = |\{a \in \mathscr{A} \,|\, L_i(a) = \text{in}\}|$ as the number of arguments citizen i accepts (and analogously $\text{out}_L(\mathscr{A})$ for rejected arguments). Moreover, we define the collection of all individual opinions as a labelling profile $\mathscr{L}$, which contains all the individual argument labellings.

Subsequently, we compute the aggregated support of arguments –and in particular, of the proposal– by means of an aggregation function (AF) that exploits the argument relationships to combine and propagate argument opinions.

Firstly, given an argument $a \in \mathscr{A}$ and a labelling L, we consider its defending and attacking arguments as $D(a) = \{b \in \mathscr{A} \,|\, b \vdash a\}$ and $A(a) = \{c \in \mathscr{A} \,|\, c \mapsto a\}$ and define:

- Positive support of a by labelling L: $Pro_L(a) = in_L(D(a)) + out_L(A(a))$ adds the number of accepted defending arguments and rejected attacking arguments.
- Negative support of a by L: $Con_L(a) = in_L(A(a)) + out_L(D(a))$ represents the number of accepted attacking arguments plus rejected defending arguments.

Next, given an argument $a \in \mathscr{A}$ and a labelling profile $\mathscr{L}$, we compute both its Indirect Opinion (IO) and its Direct Opinion (DO) by considering the labels attached to the arguments a is related with:

$$IO(\mathscr{L})(a) = \begin{cases} 1 & \text{if } Pro_{AF(\mathscr{L})}(a) > Con_{AF(\mathscr{L})}(a) \\ 0 & \text{if } Pro_{AF(\mathscr{L})}(a) = Con_{AF(\mathscr{L})}(a) \\ -1 & \text{if } Pro_{AF(\mathscr{L})}(a) < Con_{AF(\mathscr{L})}(a) \end{cases} \quad DO(\mathscr{L})(a) = \begin{cases} 1 & \text{if } in_{\mathscr{L}}(a) > out_{\mathscr{L}}(a) \\ 0 & \text{if } in_{\mathscr{L}}(a) = out_{\mathscr{L}}(a) \\ -1 & \text{if } in_{\mathscr{L}}(a) < out_{\mathscr{L}}(a) \end{cases}$$

Finally, we asses the aggregated label of an argument a by applying the following aggregation function AF (see BF in [9]) which balances both direct and indirect support:

$$AF(\mathscr{L})(a) = \begin{cases} in & \text{if } IO(\mathscr{L})(a) + DO(\mathscr{L})(a) > 0 \\ out, & \text{if } IO(\mathscr{L})(a) + DO(\mathscr{L})(a) < 0 \\ undec, & \text{if } IO(\mathscr{L})(a) + DO(\mathscr{L})(a) = 0 \end{cases} \tag{1}$$

3. Hybrid aggregation for Decidim

The two methods hereby presented tackle the aggregation of proposal support in alternative ways. On the one hand, PAM relies heavily on the quantitative (real-number) opinions that citizens express on arguments and filter out non-relevant arguments. On the other hand, TODF focuses on the dialogue structure (i.e., the attacking and defending relationships among arguments) and operates with qualitative (labelling) opinions. Therefore, we propose here a hybrid approach combining the strengths of both aggregation methods.

Inspired by the TODF aggregation function, we meant our hybrid aggregation method to grant the same importance to both direct and indirect opinions. However, Decidim just gathers positive supports for proposals to avoid possible harmful "hater" behaviours. Therefore, our hybrid aggregation function just considers indirect opinions over the proposal (i.e., those over the arguments issued when debating the proposal).

We resort to the PAM method to compute such indirect opinion so we benefit from a real-value spectrum and the possibility of discarding non-relevant arguments. However, PAM just considers arguments in favour and against the proposal, thus disregarding the tree structure of the debates in Decidim. In order to overcome this limitation, we first use the aggregation function *AF* in Eq. 1 from TODF to aggregate the opinions of every first-level subtree in the debate structure. These subtrees correspond to those trees whose roots correspond to the first-level arguments (i.e., arguments that are directly related – either by an attack or a defence relation– to the proposal). For instance, in Figure 2 c) there are 12 of such subtrees, although just 4 of them have more than a single argument.

Subsequently, the resulting aggregated qualitative opinions obtained for each first-level argument are transformed into real values by mapping `in` to 1, `out` to -1, and `undec` to 0. Finally, we are ready to use PAM to compute the final aggregated opinion on the proposal.

3.1. Hybrid aggregation results

We apply our hybrid method to real debates from Decidim Barcelona [3]. In previous work [16] we conducted a comparison of 910 proposals to assess the coincidences in support aggregation between PAM, TODF, and an base-line average computation. Results showed that they behave similarly, as about 95% match those of the average computation. Differences in the remaining c.a. 5% mostly come from: their ability to overcome average's tendency to centralise scores; PAM's aggregating quantitative opinions and deeming some proposals to be not evaluable due to lack of relevant information; and TODF relating qualitative opinions.

In order to provide some intuition of how our hybrid aggregation method[3] behaves, Table 1 explores 5 different proposals from [3]: the example proposal 50 from Figure 2 and 4 additional proposals. They have been chosen because they had high direct support (first row) and a significant number of arguments. Rows 2, 3, and 4 correspond to the number of arguments in favour, neutral, and against respectively that are first-level (in Decidim, all lower-level arguments are neutral and also added in row 5)[4]. We processed

---

[3]Source code by Marc Fernàndez is publicly available at https://github.com/marcFernandez/TODF-Argumentation/blob/tfg/NormArgumentMap.java.

[4]For instance, proposal 262 has a total of 111 arguments: 63 first-level and 52 (neutral) lower-level.

	proposal 50	prop. 179	prop. 262	prop. 1219	prop. 1256
direct support	57	446	705	141	1720
arg. in favour	11, **11**	51, **31**	31, **16**	10, **8**	49, **55**
neutral arg.	0, **0**	3, **0**	10, **0**	6, **0**	13, **0**
arg. against	1, **1**	0, **0**	22, **4**	14, **10**	8, **0**
total arg.	22, 12, **12**	55, 54, **31**	111, 63, **20**	68, 30, **18**	108, 70, **55**
Average	0.575	1.0	0.745	-0.380	0.945
PAM	0.715	0.995	0.885	-0.715	0.950
TODF	1 (in)	1 (in)	1 (in)	-1 (out)	1 (in)
Hybrid	0.75	0.98	0.9	-0.81	0.93

Table 1. Description of five different proposals from [3] and their corresponding aggregated support values as computed by Average, PAM, TODF, and Hybrid aggregation methods.

these arguments –to discard those without positive opinions or to realign those mistakenly defined– and show in bold the arguments actually used in the hybrid aggregation (see rows 2-5, where 5 adds previous rows 2-4).

Subsequent rows (from 6 to 9) in Table 1 provide the support computed by the different aggregation methods[5]. Overall, all methods coincide in the sign of the proposal aggregated support, although average shows its tendency to centrality. This is most noticeable in proposal number 1219, which is about "Limiting the invasion of dogs in public spaces". Average shows a much more bland negative value (-0.380) than PAM (-0.715), TODF (-1), or Hybrid (-0.81), which clearly result in rejecting the proposal. PAM discards most of the arguments and, since arguments against are much stronger, it rejects the proposal. As for TODF, it accepts (i.e, labels as in) more arguments against than in favour, so that it also rejects the proposal. This proposal 1219 also exemplifies well the tendency of our new hybrid method to result in values that are close to PAM and TODF values. In fact, this hybrid method results in a useful combination that inherits the advantages of each of the strenghts of PAM and TODF methods (i.e., PAM's real values allow to discard arguments that do not contribute much to the debate, and TODF's qualitative process exploits the debate structure).

4. Resilience

Debate quality has been studied based on different perspectives such as the structure of the debate (e.g., when measuring the depth of the thread of a debate [2]), the quantity of arguments issued, or number of gathered opinions. The debates we consider here are associated to decision making –that is, they are meant to decide upon the acceptance or rejection of a proposal– and thus, we propose *resilience* as a measure related to the quality (or robustness) of this decision. It is meant to perform a *sensitivity analysis* to determine how the decision is affected based on changes on the arguments and opinions. Specifically, the changes we consider are the addition of new counter-arguments and opposed opinions. In this manner, if a proposal is currently accepted, the resilience will measure the proportion of arguments (and opinions) against it that are required to be added so it becomes rejected (or conversely, if it is rejected, the changes required to

---

[5]For the sake of comparison, all values have been normalised to the interval [-1, 1]. Therefore, the ones for TODF can only take extreme values of -1, 0, or 1, whereas the rest of methods are more fine grained.

become accepted). Thus, the more changes needed, the more resilient the outcome of the debate (i.e., the decision made). In this manner, resilience can provide information to participants on how far the debate is (or what would it take to) reverse the current decision. Moreover, this would help participants that do not agree with the outcome to picture why their option is not chosen and to stimulate them to look for further participation (in terms of additional arguments or opinions).

In this context, we will refer to *aligned arguments* as those that are in line with the outcome of the debate. That is: if the proposal gets accepted, aligned arguments correspond to those in favour of the proposal. Otherwise, if the proposal is rejected, aligned arguments are those against it. Following TODF notation:

$$aligned\_arg(p) = \begin{cases} \{a \in \mathscr{A} \,|\, a \vdash p\} & \text{if } AF(p) = \texttt{in} \\ \{a \in \mathscr{A} \,|\, a \mapsto p\} & \text{if } AF(p) = \texttt{out} \end{cases}$$

Similarly, we will refer to *contrary arguments* are those that are opposite:

$$contrary\_arg(p) = \begin{cases} \{a \in \mathscr{A} \,|\, a \mapsto p\} & \text{if } AF(p) = \texttt{in} \\ \{a \in \mathscr{A} \,|\, a \vdash p\} & \text{if } AF(p) = \texttt{out} \end{cases}$$

Our resilience computes the sensitivity analysis by considering different types of argument and opinion additions. In particular, we define 4 debate extension strategies:

1. The addition of weak synthetic contrary arguments. By weak we mean that are minimally relevant (i.e., with enough opinions to be considered relevant as defined in Sect. 2.1). It measures the proportion of weak arguments that would be required to change the decision (i.e., how resilient it is w.r.t. new weak counter-arguments).
2. The addition of strong synthetic contrary arguments. By strong we mean having as many opinions as the aligned argument with the maximum number of opinions. Thus, this computes the sensitivity of the decision to strong arguments.
3. The addition of synthetic contrary arguments that are equivalent to, in terms of number of opinions, the best existing contrary argument. We only compute this dimension in case such contrary argument exists and it is relevant.
4. The addition of synthetic negative opinions about existing aligned arguments.

From these 4 debate extension strategies we define the so-called *resilience profile* as a vector $\vec{\rho}$ of four components, where each component reflects the proportion of additions that need to be included into a debate for it to change its outcome:

$$\vec{\rho}(p) = \langle \rho_1(p), \rho_2(p), \rho_3(p), \rho_4(p) \rangle$$

We compute resilience through an iterative process that consists on performing unitary synthetic additions to the debate and computing the resulting proposal support until this support changes (i.e., from accepted to rejected or vice-versa). Finally, we compute each resilience component $i \in [1,4]$ as the percentage of required synthetic additions (syn_add_i) with respect to the total elements ($total_i$) under consideration (i.e., total number of arguments for ρ_1, ρ_2, ρ_3 or total number of opinions for ρ_4):

$$\rho_i = min(100, \frac{syn\_add_i * 100}{total_i}) \tag{2}$$

4.1. Resilience results

This section is devoted to illustrate how our resilience measures the sensitivity of the decision made in a debate. Table 2 details the resilience profile for the proposals in Table 1. Furthermore, Figure 3 depicts the radar diagrams for proposals 50, 262, and 1219 (those for proposals 179 and 1256 are similar to 262 and 1219 respectively). Columns in Table 2 indicate the value (specified as a percentage) for each resilience dimension together with the values of syn_add_i and $total_i$ used in its computation. In particular, $total_1 \ldots total_3$ correspond to the number of arguments specified in 5th row of Table 1, whereas $total_4$ specifies the total number of opinions issued for those arguments.

prop. p	$\rho_1(p)$ $(syn\_add_1, total_1)$	$\rho_2(p)$	$\rho_3(p)$	$\rho_4(p)(syn\_add_4, total_4)$
50	83,33% (10, 12)	91.67% (11, 12)	100% (0, 12)	100% (18 No-Comp, 30)
179	100% (33,31)	38.7% (12,31)	100% (0, 31)	100% (114 No-Comp, 184)
262	80% (16, 20)	55% (11, 20)	100% (0, 20)	100% (79 No-Comp, 120)
1219	33.3% (6,18)	16.7% (3,18)	100% (435,18)	100% (31 No-Comp, 75)
1256	27.3% (15,55)	20% (11,55)	100% (0, 55)	100% (38 No-Comp, 275)

Table 2. Resilience profiles for the five proposals from Table 1.

When applying Equation 2, different proposals result in different values of ρ_1, varying from the 100% for prop. 179 (because becomes rejected only when adding 33 out of 31 weak synthetic arguments against it) down to the 27.3% (15 out of 55) for prop. 1256. We can also observe that the percentages for ρ_2 tend to be smaller than ρ_1, since fewer contrary arguments are required to change the decision if they are strong[6] (recall that arguments added in ρ_1 are the weakest possible). As for ρ_3, most contrary arguments turn out to be irrelevant and, thus, ρ_3 becomes the maximum possible (i.e., 100%)[7]. Finally, as for ρ_4, it turns out to be the case that for all proposals in Table 2, the iterative process goes as far as adding a number of negative opinions about aligned arguments that causes the proposal support to become non-computable (see Subsection 2.1). Again, we also interpret this situation as a maximum resilience and assign $\rho_4(p) = 100\%$.

Overall, we argue that these proposal examples (and in particular, prop. 1219 and 1256) illustrate well the need for the resilience measure, since a debate outcome that may seem at first sight quite clear (recall, from last 3 rows in Table 1, that our support aggregation methods computed values above 0.7 in absolute value) could in some cases be in fact prone to change if relatively few new contrary arguments were added.

5. Conclusions and Future Work

This paper proposes a new support aggregation method meant to be used by the Decidim participation platform. This method is derived from the combination PAM and TODF, two complementary aggregation methods previously introduced in the literature. This hybrid method inherits: from PAM, the fine-grained real values that allow to discard irrelevant arguments in the debate; and from TODF's, the qualitative process that exploits

---

[6]E.g., in 179 the strongest argument has 15 positive opinions and none negative ones.

[7]The exception being prop. 1219, where we need to add as much as 435 synthetic arguments because the best contrary argument is rather controversial (it has 4 opinions in favour and 3 against).

Figure 3. Radar diagrams of the resilience profiles for proposals 50, 262, and 1219.

the debate structure. Additionally, we propose a resilience metric for measuring the quality of the aggregated opinion and illustrate its application to debates conducted in the Decidim participatory platform. As future work, we plan to integrate both the hybrid aggregation method and the resilience measures in the Decidim open source platform. The literature on consensus in aggregation functions also deserves future exploration.

References

[1]   City of Reykjavík participation portal. http://reykjavik.is/en/participation, 2021. last visited.

[2]   Pablo Aragón, Vicenç Gómez, and Andreas Kaltenbrunner. To thread or not to thread: The impact of conversation threading on online discussion. In *11th Int. Conference on Web and Social Media*, 2017.

[3]   Decidim Barcelona. Argument data on debates in the municipal action plan participation process. https://github.com/elaragon/metadecidim/tree/master/comments, 2016.

[4]   Decidim Barcelona. City of barcelona participation portal. https://decidim.barcelona, 2021. last visited.

[5]   Consul. Free software for citizen participation. http://consulproject.org/en/, 2021. last visited.

[6]   Decidim. Participatory democracy for cities and organizations. https://decidim.org/, 2021. last visited.

[7]   Charles DeTar. *InterTwinkles: Online Tools for Non-Hierarchical Consensus-Oriented Decision Making*. PhD thesis, Media Arts and Sciences at the Massachusetts Institute of Technology, 2013.

[8]   Parlement et Citoyens. Institutional participation portal for the discussion of french laws. https://parlement-et-citoyens.fr/, 2019.

[9]   Jordi Ganzer-Ripoll, Natalia Criado, Maite Lopez-Sanchez, Simon Parsons, and Juan A. Rodriguez-Aguilar. Combining social choice theory and argumentation: Enabling collective decision making. *Group Decision and Negotiation*, 28(1):127–173, Feb 2019.

[10] Spaceship Media. Spaceship media: journalism to bridge divides. https://spaceshipmedia.org, 2016. last visited: May 2021.

[11] Participa Gencat portal. https://participa.gencat.cat/, 2021. last visited.

[12] Carole Pateman. *Rousseau, John Stuart Mill and G. D. H. Cole: a participatory theory of democracy*, page 22–44. Cambridge University Press, 1970.

[13] Petitions. UK Government and Parliament. https://petition.parliament.uk/, 2021. last visited.

[14] Juan A Rodriguez-Aguilar, Marc Serramia, and Maite Lopez-Sanchez. Aggregation operators to support collective reasoning. In *Modeling Decisions for Artificial Intelligence*, pages 3–14. Springer, 2016.

[15] Rousseau platform. https://rousseau.movimento5stelle.it/, 2021. last visited.

[16] Marc Serramia, Jordi Ganzer, Maite López-Sánchez, Juan A. Rodríguez-Aguilar, Natalia Criado, Simon Parsons, Patricio Escobar, and Marc Fernández. *Citizen Support Aggregation Methods for Participatory Platforms*, volume 319, page 9–18. IOS Press, 2019.

[17] Frank J Sorauf. The public interest reconsidered. *The Journal of Politics*, 19(4):616–639, 1957.

[18] Vicenç Torra and Yasuo Narukawa. *Modeling decisions: information fusion and aggregation operators*. Springer Science & Business Media, 2007.

[19] Vishanth Weerakkody and Christopher G Reddick. *Public sector transformation through e-government: experiences from Europe and North America*. Routledge, 2012.