October 2019

Citizen support aggregation methods for participatory platforms*

Marc SERRAMIA [a,b,1], Jordi GANZER [c], Maite LOPEZ-SANCHEZ [b],
Juan A. RODRÍGUEZ-AGUILAR [a], Natalia CRIADO [c], Simon PARSONS [c],
Patricio ESCOBAR [b], and Marc FERNÁNDEZ [b]
[a] *Artificial Intelligence Research Institute (IIIA-CSIC)*
[b] *Universitat de Barcelona (UB)*
[c] *King's College London*

Abstract. In the context of Digital Democracy, online participation platforms have emerged as innovative tools that enable citizens to participate in the decision making of their nation, region, or local government. Users can issue proposals and arguments in favour or against them and they can also support other people's arguments. This paper proposes two alternative support aggregation methods and applies them into debates conducted in the Decidim platform.

Keywords. e-governance, participation platform, debate, argumentation, information fusion, AI application.

1. Introduction

The periodical election model of representative democracy causes many citizens to feel disconnected from their governments. However, current Internet technologies provide the opportunity to bridge this gap by enabling models of on-line participatory democracy [23,11]. With this aim, participatory portals are designed to enable informed and reasoned decisions, where citizens can share their opinions with their governments. Indeed, we can find several e-participation and e-governance ICT systems such as Loomio [18], Consider.it [8], or Baoqu [4]. From these, we highlight Consul [9], which has been adopted by 100 institutions in 33 different countries and has been used by 90 million citizens. Additionally, some governments provide their own participation portals. For instance, since its launching, the UK's portal [20] has received more than 20000 petitions, some of them being extremely popular (at the time of writing this article, the proposal "Revoke Article 50 and remain in the EU" has received more than 6 million signatures). France [12] and New Zealand [19] are also making an attempt to close the gap between their parliaments and their people. Furthermore, these attempts are done at a local level, with city councils, such as Reykjavik [1] and Barcelona [7] being committed to enable participation, giving the citizens the chance to present and debate their ideas.

This paper focuses on how debates are articulated in the Decidim Barcelona [7] website. Figure 1 shows an extract of a debate about establishing free entrance to Park Güell.

[1]Corresponding Author: Marc Serramia; E-mail: marcserr@iiia.csic.es
*Research supported by the Laboratori d'Innovació Democràtica and funded by the Ajuntament de Barcelona, the Fundació Solidaritat UB and projects 2017 SGR 172, 2017 SGR 341, and AI4EU (H2020-825619).

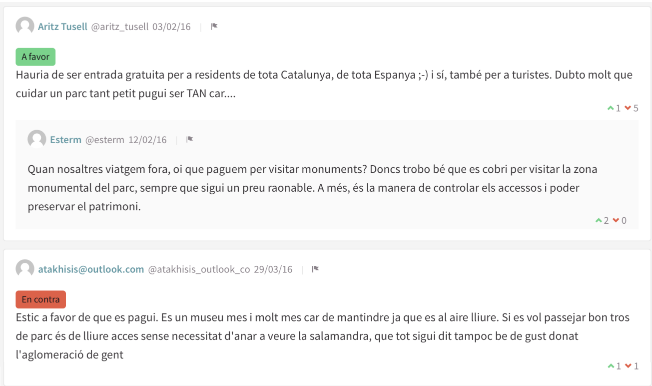


Figure 1. Part of a hierarchical Decidim debate about free entrance to Park Güell. The first argument, which is in favour, denoted by a green label, has a neutral argument response. The third argument, with a red label, is against.

Users can give arguments in favour, against, or neutral to the proposal. They can also support other people's arguments or/and respond to them with lower level arguments. The numbers near the small up and down arrows on the right of Figure 1 indicate how many citizens expressed their opinions supporting (i.e., liking) or not (disliking) each argument. This example has 2 first-level arguments: the first (green) one in favour with 1 like and 5 dislikes; the last (red) one against with 1 like and 1 dislike. By design, lower level arguments are always neutral but citizens can still indicate opinions about them. The one in Figure 1 has 2 likes and 0 dislikes.

Unfortunately, just about 10% of the proposals actually stimulate debate. To tackle this issue, we propose to aggregate (and display) information on support for proposals, so that it can serve as an stimulus for debate. Our view is that if it happens to be the case that this aggregated support is not aligned with a citizen's personal opinion, then they will be more inclined to read why it is so (i.e., to read the posted arguments) and to express their own opinions to counteract current scores. This view is aligned with the findings on counter-argumentation in [2]. In particular, this paper contributes with the adaptation and comparison of two different aggregation approaches: PAM (Proposal Argument Map) and TODF (Target Oriented Discussion Framework). PAM is an adaptation of the norm argument map [21], which applies different information fusion operators to aggregate citizen supports into a numerical value. TODF is based on argumentation, particularly we use the method in [14], which focuses on argument relations and labelled opinions on these arguments, to validate or refuting the targeted proposal. We compare these alternative methods by analysing how they aggregate the support provided in real debates from Decidim Barcelona [7]. Although they perform similarly for most cases, we illustrate the differences that emerge from focusing on fusing quantitative opinions or relating qualitative opinions. This work has been conducted in the context of the `Decidim Intel.ligent` research project, performed in collaboration with the Decidim citizen participation platform [10], with the aim of increasing citizens' involvement in debates.

We structure the paper as follows: Section 2 briefly introduces related research. Subsequent sections 3 and 4 are devoted to describe and formalise PAM and TODF respectively. These approaches are then compared in Section 5. Finally, Section 6 concludes the paper and discusses possible future paths for research.

2. Related work

We are witnessing new forms of participatory and deliberative democracy based on computer mediated communication [15]. UK's portal [20] enables participatory but not deliberative democracy, since proposals are not discussed. Your Priorities [16] structures debates in two columns, similarly to PAM (see Section 3), grouping arguments in favour or against the proposal to facilitate decision making. Although it sorts arguments by their number of supports, it does not aggregate supports. Alternatively, Consul [9] displays threaded discussions that relate arguments as (neutral) replies. As explained in Section 4, TODF relates arguments as attacks or defences, which allows us to aggregate supports in a meaningful way.

On-line petition platforms are attracting a lot of attention from the research community [15,23,11]. However, how to effectively introduce discussions remains as an open challenge [2]. Artificial Intelligence techniques have been proposed for tackling this issue. For example, the work in [13] propose the application of information fusion and optimisation techniques to collective decision support making. More importantly, Klein [17] takes a large scale argumentation approach to facilitate deliberation. In fact, PAM is inspired in his argument maps. Finally, argumentation is formally considered for judgement aggregation in [3]. TODF can be somehow seen as an extension of this, since in addition to attack relations it also considers defence relations to better fit on-line debates.

3. Proposal Argument Map (PAM)

Before formalising the aggregation methods, we introduce some basic concepts. Thus, we consider a proposal as a suggestion put forward for consideration by others and an *argument* as a statement providing a reason in favour of or against it. An *opinion* is a (quantitative or qualitative) value that someone assigns to an argument in order to express their support (or lack of).

Next, in the vein of [21], we define a proposal argument map as follows:

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).*

Specifically, for a given argument $a \in A_p$, $\kappa(a) = 1$ (if a is in favour of p), $\kappa(a) = 0$ (if a is neutral), and $\kappa(a) = -1$ (if a is against p). Hence, κ allows us to group arguments in favour of and against the proposal in two separate sets (A_p^+ and A_p^- respectively) and to disregard neutral arguments since they do not contribute much to the final decision. Figure 2 (left) shows an example of a proposal argument map (PAM), where A_p^+ and A_p^- are displayed column-wise and citizens can express their opinion about individual arguments with a 5-star scale interface. This scale in fact corresponds to an Opinion Spectrum $[lb, ub]$: a real interval where lb stands for the lower bound and means the most negative opinion, and ub, the upper bound, is used to cast the most favourable opinion.

Besides the 5-star scale, each argument depicts the number of citizens that have provided their opinion on it. Formally, we consider an argument $a \in A_p$ to be 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. Although we

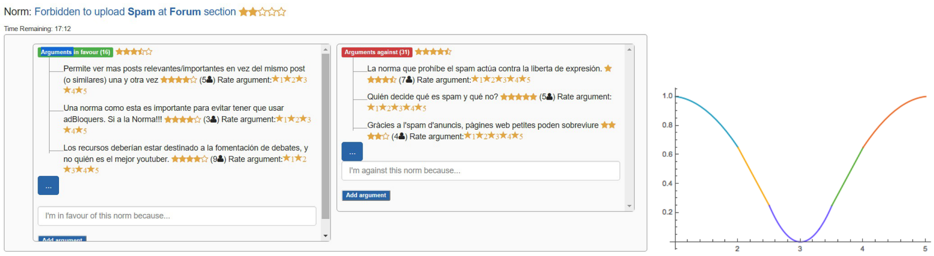


Figure 2. Left: Example of a proposal argument map (PAM) displaying the support scores for: the proposal; the set of arguments in favour, A_p^+; the set of arguments against, A_p^-; and individual arguments. Citizens affect the aggregated opinions by manually scoring individual arguments. Right: A (piece-wise) importance function.

could average opinions, some inconveniences would arise. Firstly, averages of polarised opinions will lead to central (neutral) scores no-one would have expressed. Secondly, a majority of neutral opinions would mask strong minority opinions. Thus, in what follows, we advocate for alternative aggregation operators to be used, not only to assess argument support, but also for sets of arguments as well as for the proposal.

3.1. Argument support

When considering issued opinions, we argue that some opinions may be more relevant than others since, for example, neutral opinions can be associated with indecision, whereas extreme opinions may signal strong (clear) positions. Thus, we consider an opinion importance function I (see [21]) to weight citizen opinions in a weighted mean (WM_w). The right-hand side of Figure 2 shows a possible importance function, which associates the highest weights to extreme opinions and zero weight to neutral ones.

Formally, given an argument $a_i = (s_i, O_{a_i})$ having the set O_{a_i} of issued opinions, we define the argument support as a weighted mean:

$$S_{arg}(a_i) = WM_w(O_{a_i})$$

with weights $\left(\frac{I(o^i_1)}{T}, \ldots, \frac{I(o^i_{n_i})}{T}\right)$, where $T = \sum_{j=1}^{n_i} I(o^i_j)$ corresponds to their normalisation factor; $o^i_j \in O_{a_i}$ stands for the support opinion provided by citizen j over argument a_i; $I(o^i_j)$ accounts for its corresponding importance; and $n_i = |O_{a_i}|$ is the number of citizens that have issued an opinion on argument $a_i \in A_p$.

Additionally, we propose a notion to determine the relevance of the supporting opinions, and thus be able to discard weak (non-relevant) arguments that would hinder the aggregated support. Formally, we consider an argument a_i to be α-*relevant* if it has a significant number of opinions (i.e., if $\forall a_j \in A,\ |O_{a_i}| \geq \alpha \cdot |O_{a_j}|$, where $\alpha \in [0,1]$) and it has a enough support (i.e., if $S_{arg}(a_i) > \frac{lb+ub}{2}$).

3.2. Argument set support

In order to scale up the computation of the support to a set of arguments A, we also use an aggregation operator. First, from a set of arguments A, we consider the set $R_\alpha(A)$ of its k' α-*relevant* arguments ($k' = |R_\alpha(A)|$) and assess the support of this argument set A by using the WOWA (Weighted Ordered Weighted Average) operator [22]:

$$S_{set}(A) = \begin{cases} WOWA_{w,q}(S_{arg}(a_1), \ldots, S_{arg}(a_{k'})) & \text{if } R_\alpha(A) \neq \emptyset \\ \text{non-computable} & \text{otherwise} \end{cases}$$

The parameters [22] w, τ, q, and T in WOWA take the values:

$$w = \left(\frac{\sum_{j=1}^{|O_{a_1}|} I(o_j^1)}{\tau}, \ldots, \frac{\sum_{j=1}^{|O_{a_{k'}}|} I(o_j^{k'})}{\tau}\right), \quad \tau = \sum_{i=1}^{k'} \left(\sum_{j=1}^{|O_{a_i}|} I(o_j^i)\right), \quad o_j^i \in O_{a_i}$$
$$q = \left(\frac{I(S_{arg}(a_{\sigma(1)}))}{T}, \ldots, \frac{I(S_{arg}(a_{\sigma(k')}))}{T}\right), T = \sum_{i=1}^{k'} I(S_{arg}(a_{\sigma(i)})), \quad a_{\sigma(i)} \in R_\alpha(A)$$

Briefly, w and q are weighting vectors normalised by τ and T respectively, $I(o_j^i)$ is the importance of the opinion issued by citizen j over an α-relevant argument a_i, and $a_{\sigma(i)}$ is the α-*relevant* argument with the i^{th} largest support. Intuitively, w weights "participation importance" as the sum of the opinion importances so that arguments with more participation and stronger opinions count more. As for q, it weights the importance of the numbers being aggregated, in this case the importance of each argument's support (please refer to [21] for a detailed explanation). Notice that we only compute the aggregated support of an argument set whenever there are α-*relevant* arguments, otherwise we consider we lack enough quality opinions.

3.3. Proposal support

Finally, given a proposal p and the (non-empty) set of α-*relevant* arguments, we define $R_\alpha(A_p^+)$ as the set of α-*relevant* arguments in favour of p and $R_\alpha(A_p^-)$ as the set of arguments against it, and combine their respective supports applying the same WOWA operator. Specifically, we consider $|R_\alpha(A_p^+)| = k_1$, $|R_\alpha(A_p^-)| = k_2$ and $k_1 + k_2 > 0$ (since $R_\alpha(A) \neq \emptyset$) and compute the proposal support as:

$$S_{prop}(p) = WOWA_{w,q}\big(S_{set}(R_\alpha(A_p^+)), ub + lb - S_{set}(R_\alpha(A_p^-))\big)$$

$$w = \left(\frac{\sum_{i=1}^{|R_\alpha(A^+)|}(\sum_{j=1}^{|O_{a_i}|} I(o_j^i))}{\tau}, \frac{\sum_{i=1}^{|R_\alpha(A^-)|}(\sum_{j=1}^{|O_{\overline{a}_i}|} I(\overline{o}_j^i))}{\tau}\right), \; T = I(S_{set}(A_p^+)) + I(ub + lb - S_{set}(A_p^-)),$$
$$\tau = \sum_{i=1}^{|R_\alpha(A^+)|}(\sum_{j=1}^{|O_{a_i}|} I(o_j^i)) + \sum_{i=1}^{|R_\alpha(A^-)|}(\sum_{j=1}^{|O_{\overline{a}_i}|} I(\overline{o}_j^i)), \; q = \left(\frac{I(S_{set}(A_p^+))}{T}, \frac{I(ub+lb-S_{set}(A_p^-))}{T}\right),$$

where $o_j^i \in O_{a_i}$ is the j^{th} opinion of an argument $a_i \in R_\alpha(A_p^+)$ in favour of p , and $\overline{o}_j^i \in O_{\overline{a}_i}$ is the j^{th} opinion over an argument $\overline{a}_i \in R_\alpha(A_p^-)$ against p (details can be found in [21]).

4. Target oriented discussion framework (TODF)

Here we propose the Target Oriented Discussion Framework (TODF) as 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 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 3a depicts an example of a TODF, where arguments a_1 and a_2 attack the target ($a_1 \mapsto \tau$, $a_2 \mapsto \tau$) and arguments a_3 and a_2 defend the target and a_1 respectively ($a_3 \vdash \tau$, $a_2 \vdash a_1$). Moreover, as detailed in [14], TODFs have to fulfil several properties (such that all arguments are descendants of the target or that there are no cycles in their relationships nor reflexivity).

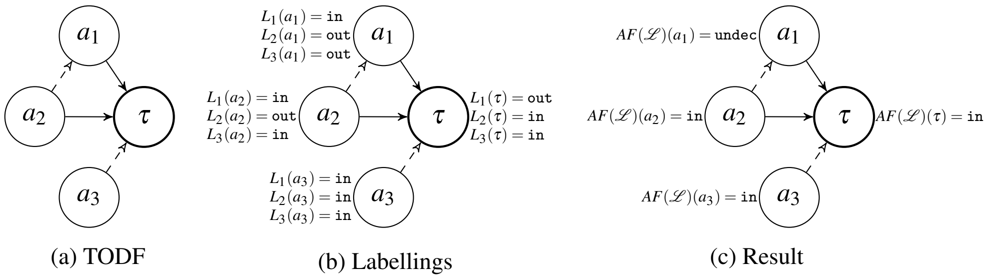


Figure 3. An example target oriented discussion framework. Specification of: (a) argument and target relationships ($a_1 \mapsto \tau, a_2 \mapsto \tau, a_2 \vdash a_1, a_3 \vdash \tau$); (b) 3 individual opinions (labellings L_1, L_2, and L_3); and (b) aggregated opinion ($AF(\mathscr{L})$).

Figure 3b illustrates that citizens express their opinions by assigning labels to arguments and the proposal (which is a particular argument) in our Target Oriented Discussion Framework. Such labels are: *in*, meaning they agree with the argument; *out*, if they disagree; and *undec*, if they are neutral or not sure. In fact, *undec* is the default label, so that it is assigned whenever a citizen does not cast an opinion on an argument. Formally, each citizen cz provides an argument labelling, which is a function $L_{cz} : \mathscr{A} \to \{in, out, undec\}$. Then, given an argument labelling L and a subset $A \subset \mathscr{A}$ of arguments, we denote the number of arguments *accepted* in A by L as $\mathtt{in}_L(A) = |\{b \in A \mid L(b) = \mathtt{in}\}|$ and the number of *rejected* arguments as $\mathtt{out}_L(A) = |\{b \in A \mid L(b) = \mathtt{out}\}|$. Moreover, given a population of participant citizens $\{cz_1, \ldots, cz_n\}$, we consider a labelling profile which is a tuple containing all their argument labellings $\mathscr{L} = (L_{cz_1}, \ldots, L_{cz_n})$.

From the labelling profile (i.e., the collection of all individual opinions), we can then compute an aggregated support (see Figure 3c) of arguments, and in particular, of the proposal. With this aim, we propose the usage 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 $A(a) = \{c \in \mathscr{A} | c \mapsto a\}$ respectively and define its positive and negative support as:

- $Pro_L(a)$, positive support of a: stands for the number of accepted defending arguments and rejected attacking arguments by labelling L.
  Formally: $Pro_L(a) = in_L(D(a)) + out_L(A(a))$.
- $Con_L(a)$, negative support of a: represents the number of accepted attacking arguments and rejected defending arguments by L.
  Formally: $Con_L(a) = in_L(A(a)) + out_L(D(a))$.

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

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

and we denote its Direct Opinion (DO) as:

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

Finally, we asses the aggregated label of an argument a by applying the following aggregation function AF (see BF in [14]) which balances both direct and indirect support:

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

It is worth mentioning that we assume the graph associated to the TODF is a DAG. This allows to identify leaf arguments (i.e., those without any other arguments attacking nor defending them) and to compute aggregated support by evaluating them first.

5. Analysis and comparison of aggregation methods

Following the descriptions of the Proposal Argument Map (PAM) and Target Oriented Discussion Framework (TODF) methods in previous sections, this section analyses how these models can aggregate the support provided in real debates about Barcelona's municipal action plan [6]. These debates were conducted on-line at the Decidim Barcelona website [7], and the data is publicly available at [5].

Figures 1, 2, and 3 illustrate that PAM and TODF methods consider dialog structures that differ from that actually used by Decidim. Therefore, PAM and TODF need to be adapted. For PAM, arguments are grouped in two sets: A_p^+ containing arguments issued in favour of the proposal and those posed in response to arguments against the proposal; and A_p^- containing arguments issued against the proposal and those posed in response to arguments in favour of the proposal. As for TODF, defence $a \vdash a'$ relationships are established for arguments a created in favour of other arguments a', and attack $a \mapsto a'$ relationships are established whenever an argument a is explicitly against the target ($a' = \tau$) or it is simply created as a response to any other argument $a' \in \mathscr{A}$. As already mentioned, by design, arguments attacking or defending arguments other than the proposal are always set to be neutral in Decidim. However, we instead assume arguments are issued to revoke previous arguments because this is what appears to be most often the case in dialog threads (the second argument in Figure 1 is an example of this). Indeed, Aragon et al. [2] analysed the same dialogue data and concluded that most conversation cascades came from counter-argumentation.

PAM, TODF, and Decidim also differ in opinion assessment. Hence, in order to use the dialogue data gathered in Decidim with PAM and TODF, and to be able to compare their results, we establish the following correspondences: First, a like in Decidim is interpreted as a 5 in PAM's opinion spectrum ($ub = 5$) and as a *in* label in TODF; Second, we map a dislike in Decidim to a 1 in PAM ($lb = 1$) and as an *out* label in TODF; and third we map an *in* in TODF to the interval [4,5] in PAM, an *undec* to (2,4), and an *out* to

[1,2]. However, for the sake of comparison, we also consider an (overlapping) extended interval where *in*=[3.5,5], *undec*=(1.5,4.5), and *out*=[1,2.5]. As for opinion aggregation, since Decidim does not aggregate support, we consider an average approach as the baseline to compare with PAM and TODF aggregation methods. Specifically, this method averages as many 5's as the number of likes in (positive) arguments in favour of the proposal plus the number of dislikes in (negative) arguments against the proposal together with as many 1's as the number of dislikes in positive arguments plus the number of likes in negative arguments:

$$Average(p) = \frac{5 \cdot (likes\_in(A_p^+) + dislikes\_in(A_p^-)) + 1 \cdot (likes\_in(A_p^-) + dislikes\_in(A_p^+))}{likes\_in(A_p) + dislikes\_in(A_p)}$$

As for PAM aggregation computation, we use the importance function from right-hand-side of Fig 2 and an $\alpha = 0.3$ for argument relevance.

Real data in Decidim [5] covers 10860 proposals and the participation of more than 40,000 citizens. Most of these proposals have no comments in their debate section, and removing the proposals without comments leaves us 5199 (47,87%) with at least one comment. Having just one comment may not be enough information to assess support for the proposal, and requiring at least a vote of some sort (like/dislike) for any comment further reduces the number of proposals to 1102 (10,15%). When we eliminate those proposals with only neutral arguments we are left with 910 (8,38 %) proposals only which have at least one non neutral argument and at least one like or dislike.

5.1. Results

We computed the support aggregation for these 910 proposal debates using each of the PAM[2], TODF[3], and average methods. Table 1 shows the number of cases in which the results for pairs of these three aggregations mechanisms match when we use the two different interval representations discussed above. Thus, for example, any aggregated value resulting from PAM or average that belongs to [4,5] or [3.5,5] matches an *in* from TODF. Thus, the expanded (overlapping) intervals help to assess meaningful differences. Of the 910 proposals, the PAM aggregation method deems 49 proposals to be not evaluable due to lack of relevant information. These account for most of the differences (5,38% out of the 5,93% for expanded opinion intervals) when comparing PAM to the average in the first row. In fact, being able to identify such cases is the main virtue of our PAM method when reducing the richness of its opinion spectrum down to like/dislike opinions.This difference also applies when comparing PAM to TODF (third row), which clearly differ in 9.45% of cases. However, focusing on fusing quantitative opinions (as for PAM) or relating qualitative opinions (TODF) lead to different aggregated opinions in 4.07% of proposals. Figure 4a illustrates well this difference. On the one hand, TODF accepts proposal 408 (see τ in green) because it accepts defending argument a2316 and cannot decide about defending arguments a570 and a5138. On the other hand, PAM discards the proposal because it just considers argument $a1944$ to be α-relevant (it has 4 likes whereas the rest of arguments have 1 like at most) and to be against the proposal. Conversely, the average outputs a 2,33, a rather neutral response that supports our claim about average tending to centralise scores.

[2] Source code available at: https://bitbucket.org/marcserr/commentevaluator

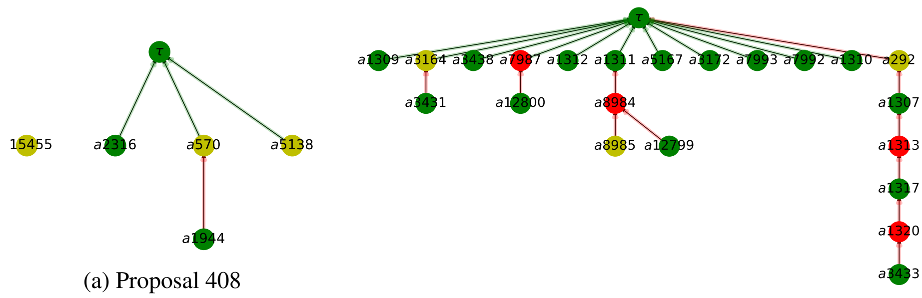
[3] Source code available at: https://github.com/marcFernandez/TODF-Argumentation

	Non-overlapping intervals	Expanded intervals
Average - PAM	790 (86,81%)	856 (94,07%)
Average - TODF	814 (89,45%)	871 (95,71%)
PAM - TODF	792 (87,03%)	824 (90,55%)

Table 1. Pairwise method comparison in number of matches (and %) for 910 proposals. Non-overlapping opinion intervals: out/undec/in = [1,2]/(2,4)/[4,5]. Expanded intervals: out/undec/in = [1,2.5]/(1.5,4.5)/[3.5,5].

The second row in Table 1 also highlights the similarities between TODF and the average computation, which agree on as many as 95.71% of proposals. We consider this to be desirable, since in general (except for the few cases aforementioned) average seems a rather natural way of combining different opinions. Furthermore, these similarities stand no matter how complex debates get. As an example, Figure 4b shows proposal 50, which has gathered 22 arguments for and against it. As before, the green τ indicates TODF accepts this proposal by: accepting 9 arguments a that defend it ($a \vdash \tau$); discarding argument a7987; and not being able to decide about arguments a3164 and a292 because they are attacked by accepted arguments a3431 and a1307 respectively (notice that in fact, a292 undecision further comes from the chain of accepted and rejected arguments that start with argument a3433). TODF's proposition acceptance clearly matches the result of the average, which corresponds to a 4.15. Additionally, PAM evaluates the proposal with an aggregated opinion of 4.43 in the [1,5] interval.

Overall, we can conclude that the two methods hereby presented tackle the computation of proposal support aggregation 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 opinions. 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, finding a hybrid approach able to combine the strengths of both aggregation methods seems promising, but still remains as future work.



(a) Proposal 408

(b) Proposal 50

Figure 4. Debate structure of proposals 50 and 408 and evaluation using the TODF approach.

6. Conclusions and future work

With the aim of promoting deliberative democracy, this paper proposes PAM and TODF, two alternative support aggregation methods to be applied in real on-line debates. Briefly, PAM focuses on fusing quantitative (numerical) support opinions whereas TODF considers qualitative opinions (labels) on related arguments. We compare them by computing the aggregated support of 910 proposals created and discussed in the Decidim Barcelona website. Results show that they behave similarly and in a rather natural way, as they mostly match those of an average computation. However, they overcome, in alternative ways, the problems raised in specific cases by the use of average. As for now, it is a decision of the platform designers which aggregation method to use, since PAM is able to manage a range of opinion values and disregards non-relevant arguments whereas TODF is more faithful to the structure of the dialogue. In future work we plan to design a hybrid approach able to combine the strengths of both methods. In the near future, though, we are working together with the Decidim team to consider alternative importance functions and to extend the platform functionality so that citizens can express a range of opinions instead of the current binary like/dislike system.

References

[1] City of Reykjavík participation portal. http://reykjavik.is/en/participation, 2019.

[2] Pablo Aragón, Andreas Kaltenbrunner, Antonio Calleja-López, Andrés Pereira, Arnau Monterde, Xabier Barandiaran, and V. Gómez. Deliberative platform design: The case study of the online discussions in decidim barcelona. In *International Conference on Social Informatics*, pages 277–287. Springer, 2017.

[3] Edmond Awad, Richard Booth, Fernando Tohmé, and Iyad Rahwan. Judgement aggregation in multi-agent argumentation. *Journal of Logic and Computation*, 27(1):227–259, 2017.

[4] Baoqu. http://beta.baoqu.org/, 5 2019. last visited.

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

[6] Decidim Barcelona. Municipal action plan participation process (catalan and spanish description). https://www.decidim.barcelona/processes/pam, 2016.

[7] Decidim Barcelona. City of barcelona participation portal. https://decidim.barcelona, 2019.

[8] Consider.it. https://consider.it/, 5 2019. last visited.

[9] Consul. Free software for citizen participation. http://consulproject.org/en/, 2019.

[10] Decidim. Free OS participatory democracy for cities and organizations. https://decidim.org/, 2019.

[11] 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.

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

[13] José Luis Fernández-Martínez, Maite Lopez-Sanchez, Juan Antonio Rodríguez-Aguilar, Dionisio Sánchez-Rubio, and Berenice Zambrano Nemegyei. Codesigning participatory tools for a new age: A proposal for combining collective and artificial intelligences. *International Journal of Public Administration in the Digital Age (IJPADA)*, 5(4):1–17, 2018.

[14] 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.

[15] Barry N Hague and Brian D Loader. *Digital democracy: Discourse and decision making in the information age*. Psychology Press, 1999.

[16] Iceland. Your priorities: edemocracy web application. https://www.yrpri.org/domain/3, 2019.

[17] Mark Klein. Enabling large-scale deliberation using attention-mediation metrics. *Computer Supported Cooperative Work (CSCW)*, 21(4-5):449–473, 2012.

[18] Loomio. Software to assist collaborative decision-making processes. https://www.loomio.org/, 2016.

[19] New Zealand's ministry of Justice consultation hub. https://consultations.justice.govt.nz/, 2019.

[20] Petitions. UK Government and Parliament. https://petition.parliament.uk/, 2019. last visited.

[21] 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.

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

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