Committee Selection using Attribute Approvals

Venkateswara Rao Kagita
National Institute of Technology
Warangal, India
venkat.kagita@nitw.ac.in

Arun K Pujari
Mahindra University
Hyderabad, India
arun.k.pujari@gmail.com

Vineet Padmanabhan
University of Hyderabad
Hyderabad, India
vcpnair73@gmail.com

Haris Aziz
UNSW Sydney
Sydney, Australia
haris.aziz@unsw.edu.au

Vikas Kumar
University of Delhi, India
Central University of Rajasthan, India
vikas007bca@gmail.com

ABSTRACT

We consider the problem of committee selection from a fixed set of candidates where each candidate has multiple quantifiable attributes. Instead of voting for a candidate, a voter is allowed to approve the preferred attributes of a given candidate. Though attribute-based preferences capture several important real-life scenarios, committee selection problem with attribute approval of voters has not properly formalized or studied. We present a detailed study of axioms, rules, algorithms, and computational complexity for the setting. Apart from extending previous axioms and rules, we design two new algorithms that are especially appropriate for our model.

KEYWORDS

multi-winner voting; committee voting; representation; social choice

ACM Reference Format:

Venkateswara Rao Kagita, Arun K Pujari, Vineet Padmanabhan, Haris Aziz, and Vikas Kumar. 2021. Committee Selection using Attribute Approvals. In *Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), Online, May 3–7, 2021*, IFAAMAS, 9 pages.

1 INTRODUCTION

Committee selection, from a set of candidates, by aggregating voters' preferences is a fundamental problem of social choice theory and has recently received considerable attention from the AI community [2–4, 32, 34]. Generally, each candidate possesses a set of quantifiable attributes that make a candidate suitable or otherwise. In this paper, we formally specify and study an expressive social choice setting that captures many real-world scenarios, such as employee hiring and parliamentary committees. In such a setting, candidates are judged across various attributes, such as educational qualification and experience. For each of these attributes, a candidate satisfies one of the attribute-values such as having an MBA. Voters are asked to express approvals over attribute values. For example, a voter could approve all education qualifications from higher education, including MBA. These approvals over attribute values indirectly count as approvals towards candidates that satisfy those values. Given the approval ballots of the voters, the objective is to select a committee W of k candidates.

Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), U. Endriss, A. Nowé, F. Dignum, A. Lomuscio (eds.), May 3–7, 2021, Online.

Attribute approvals based committee selection allow voters to express more specific opinions by allowing separate approval-ballots for different perspectives or dimensions of a candidate. It can exploit the multifaceted voter interests and select a committee that is more accurate than the committee selected using voters' approvals/scores over candidates. Attribute-based preference aggregation can provide more objectivity in preference reporting and aggregation by basing the preferences over a candidate on attributes that are approved by the voter and which are satisfied by the candidate. The two-step process adds a layer of explainability for the voters' preferences that is important in candidate selection problems, such as hiring. Furthermore, the committees formed over attribute preferences are less vulnerable to the practices, such as bribery [14, 18], and manipulation[1]. The overlap of attributes among several candidates and the non-trivial complexity of investigating these practices at the attribute level makes the system less susceptible. However, the major challenge here is to effectively exploit the voters' approval on attributes to determine the final set of candidates to be selected into the committee. Thus far, there has not been any attempt in this direction in the context of committee formation, although several other approaches for committee formation do exist.

Contributions. Our key contributions are as follows. We present a new preference aggregation setting that captures many real-life scenarios and simultaneously generalizes several important social choice settings. First, we formally present the formulation of the problem and several analytical results of committee selection with approvals on attributes. Second, in the context of this problem, we revisit existing properties desirable of the outcome committee, such as homogeneity, consistency, monotonicity, unanimity and justifiable representation, and provide the adapted definitions. We show that the two properties, unanimity and justified representation, are not satisfied by standard aggregation techniques (discussed in Section 4.1) while considering attribute based approvals. We focus on some key properties like *weak unanimity, strong unanimity, simple justified representation, compound justified representation, simple proportional justified representation*, and *compound proportional justified representation* and prove corresponding complexity results. We also show that computing a justified committee with the highest approval voting or highest satisfaction approval voting is an NP-complete problem. We propose an approximation scheme for this problem and perform worst case analysis. Third, we propose a

[1] Manipulability in a voting system is a scenario wherein a voter submits a disingenuous ballot that favors his/her interesting outcome against his true preferences [36].

new aggregation rule based on a greedy approach and show that this rule satisfies the unanimity and simple justified representation, but not the simple proportional justified representation as well as compound justified representations. We prove the later problems to be NP-complete. Finally, we introduce a new rule called Justified Approval Voting (JAV) that satisfies the simple proportional justified representation alongside other desirable properties.

The organization of the paper is as follows. In Section 2, we discuss work in the area that is closely related. In Section 3, we formulate and analyze the committee selection problem with attribute approval voting. In Section 4, we discuss adaptations of standard voting rules and properties to our setting. The detailed analysis of justified representation is presented in Section 5. We show that justified committee with the highest Approval Voting (AV) or Satisfaction Approval Voting (SAV) are NP-complete problems and propose an approximation scheme. We adapt Greedy Approval Voting (GAV) to attribute approvals in Section 6 and examine different properties. In Section 7, we propose a new rule called Justified Approval Voting (JAV) that satisfies simple justified proportional representation. Section 8 provides conclusions and scope for future work.

2 RELATED WORK

Attribute-level preferences exist in various domains such as food [17, 29], health care [1], housing [13], farming [22], airline services [15], technology product markets [35], job [21], e-transactions [12], and travel [19]. The significance of attributes in the committee formation has been highlighted by various researchers [5, 9, 25] wherein these works focus on committee formation with constraint satisfaction using voters' approval ballots on candidates. For instance, Brams et al. [5] represent a candidate with two attributes, "Region of the candidate" and "Specialty" and the constraint could be "10% of candidates in the committee should be from region A".

Another significant work on multiple attributes is by Lang et al. [24, 25] where they consider proportional representation in the committee selection, but not necessarily based on voters' approvals. Briefly, the problem considered by Lang and Skowron [24, 25] is to find a committee that closely satisfies the desired proportional distribution for each of the attributes and consequently, the candidate representation takes precedence. Whereas we address the problem of selecting k candidates given the voters' attribute approvals apriori. Therefore, the constraints in our problem setting are on the proper representation of the voters.

Furthermore, considerable research has been done in the area of voting in combinatorial domains [23, 26, 28, 37, 38] to address the problem of collective decision-making over several domains or attributes given the voters/agents conditional preferences. For example, if voters have to agree on a common menu to be composed of a main course and wine and the conditional preference of some voter could be "if the main course is meat then I prefer red wine, otherwise I prefer white wine". The works in Lang et al. [23, 26] address this problem by decomposing the problem into smaller problems and sequentially making decisions over individual domains. At each stage, their approach considers voters' conditional preferences of the current domain w.r.t. the previously selected candidates of other domains to determine a score of an individual candidate. Xia et al. [38] defined order-independent sequential composition of voting rules and study properties of different voting rules in this context and further improvements were made by works in [28, 37]. None of these schemes work for the current problem, nor can they be trivially extended.

Table 1: Summary of notation

Notation	Description
n	Number of voters
m	Number of candidates
d	Number of dimensions
k	Target committee size
n_a	Number of distinct attribute-values over all dimensions.
V	Set of all voters
C	Set of all candidates
C_i	Set of candidates approved by a voter v_i when $d = 1$
$c_i[j]$	Attribute value of a candidate c_i on dimension j
D^j	Set of domain values on domain j
C_i^j	Set of attributes approved by a voter v_i on dimension j
W^j $(resp. C^j)$	Multi-Set of attributes of W $(resp. C)$ on dimension j
W	Target committee
$V_i (resp.\ V_i'), i \in \mathbb{N}$	Set of voters from $V (resp.\ V')$ who approved c_i when $d = 1$.
$V_{c_i[j]}$	Set of voters from V who approved an attribute $c_i[j]$.
V_W	Multi-set of voters from V who approved attributes of a committee W
c_*	candidate with the highest utility score
V_*	Set of voters who approved a candidate c_*

3 ATTRIBUTE APPROVAL VOTING

Let $V = \{v_1, v_2, \ldots, v_n\}$ be the set of voters and $C = \{c_1, c_2, \ldots, c_m\}$ be the set of candidates. Each candidate $c_i, 1 \leq i \leq m$, is associated with a d-dimensional attribute vector or simply, d attributes. The attribute value $c_i[j]$ of candidate c_i on dimension j is from a domain $D^j, 1 \leq j \leq d$. Let C_i^j denote the set of values in D^j that are approved by voter v_i. The goal is to select a committee W of k candidates, given voters' approvals over attributes, i.e., $C_i^j, \forall v_i \in V, j \in [1, d]$. We use $V_{c_i[j]}$ to denote the set of voters who have approved an attribute-value $c_i[j]$. Table 1 summarizes the notation used in this paper. Next, we illustrate the setting with a toy example.

Example 1. *A house of parliament has to select a finance committee of k candidates. Let say that each candidate is a tuple of $d = 4$ attributes,* Educational Qualification (EQ), Political Experience (PE), Age, *and* the Political Party (PP) *to which he/she belongs. For example, $c_1[4] = 'A'$ in Table 2 represents the political party of the candidate c_1 to which he belongs. $D^4 = \{A, B, C\}$ is the domain of attribute* Political Party. *A set of attributes-values in a particular attribute forms the domain of that attribute. Voters approve desired attribute-values from each attribute's domain. For example, voter v_1's approvals could be as follows.*

Table 2: Toy Example

Candidate	*EQ*	*PE*	*Age*	*PP*
c_1	*MBA*	*5-10*	*Junior*	*A*
c_2	*MA*	*<5*	*Junior*	*B*
c_3	*MA*	*>20*	*Senior*	*C*
c_4	*PhD*	*10-20*	*Senior*	*B*
c_5	*BA*	*5-10*	*Senior*	*C*
c_6	*PhD*	*<5*	*Senior*	*A*

$C_1^1 = \{MBA, PhD\}$ $C_1^2 = \{> 20\}$ $C_1^3 = \{Senior\}$ $C_1^4 = \{A, B\}$.
For instance $C_1^1 = \{MBA, PhD\}$ *denotes that voter* v_1 *approves* MBA *and* PhD *from the domain* $D^1 = \{MBA, MA, BA, PhD\}$ *of attribute* Educational Qualification. *The task is to select k candidates using these voters' approvals over attribute-values.*

The model presented in this paper is more general than party-based approval studied recently [11]. In the latter model, each candidate can be viewed as having exactly one attribute which corresponds to the party the candidate belongs to. On the other hand, in our model, a candidate can have multiple attributes. Our model is also more general than approval-based committee voting [3, 20]. Approval-based committee voting can be reduced to our setting by having a unique attribute corresponding to each candidate. Voters can be seen as approving these attributes rather than candidates.

Next, we present an example that illustrates that committees selected based on attribute approvals can provide a decent representation of certain attributes that a classic setting of voters approving candidates disregards.

Example 2. *Consider the following set of candidates from which the problem is to select a committee* W *of two candidates:* $c_1 = [w_1, x_1, y_1, z_3], c_2 = [w_3, x_2, y_2, z_2], c_3 = [w_1, x_2, y_1, z_2]$, *and* $c4 = [w_2, x_1, y_2, z_1]$. *Let us take a hypothetical scenario wherein a set of* $\frac{n}{2}$ *voters (say* V'*) have the preferences* $\{\{w_1\}, \{x_1\}, \{y_1\}, \{z_1\}\}$ *for the committee* W *and rest of the* $\frac{n}{2}$ *voters (say* V''*) have* $\{\{w_2\}, \{x_2\}, \{y_2\}, \{z_2\}\}$ *as their preferences. Given a choice of approving candidates, voters from* V' *approve* c_1 *to maximize their preferred attributes for the committee* W*. Similarly, voters from* V'' *approve* c_2*. However, none of* c_1 *and* c_2 *covers attributes* w_1 *or* z_1 *desired by the committee according to voters' preferences. Whereas* c_3 *and* c_4 *together satisfy all the attributes approved by both* V' *and* V''*, who are otherwise could be selected into committee* W *if we consider attribute approvals.*

4 VOTING RULES AND PROPERTIES

There are different aggregating rules known in the context of committee selection's classic setting wherein voters approve candidates. We study these rules in the context of attribute approval based committee selection in Subsection 4.1. Subsection 4.2 reviews the desired properties of these rules in the current framework. In Subsection 4.3, we discuss one of the significant properties named *unanimity* that is nontrivial to extend to the attribute level setting.

4.1 Attribute Level Voting Rules

Approval Voting (AV) – Approval Voting selects a committee W that maximizes $\sum_{v_i \in V} |W \cap C_i|$, where C_i is the set of candidates approved by a voter v_i [6]. When voters submit their approvals on the attributes, the *approval voting score* of c with respect to V is defined as $AV(c, V) = \sum_{v_i \in V} \frac{\sum_{j=1}^{d} |\{v_i | c[j] \in C_i^j\}|}{d}$. The approval voting score of a committee W is $AV(W, V) = \sum_{c \in W} AV(c, V)$. Hence, *Approval Voting (AV)* rule selects W with highest $AV(W, V)$, which can be computed by maximizing $\sum_{v_i \in V} \sum_{j=1}^{d} |W^j \cap C_i^j|$, where W^j is the set of attributes on dimension j of W and C_i^j is voter v_i's attribute approvals on dimension j. Computing approval score for each attribute involves scanning n ballots and can be done in $O(n_a n)$ where n_a is the number of distinct attributes values over all dimensions. Approval score of a candidate can be computed in $O(md)$ and identifying top-k candidates takes $O(klog(m))$. Hence, the running time of AV for attribute-approval is $O(n_a n + md + klog(m))$.
Satisfaction Approval Voting (SAV) – SAV [7] selects a committee W that maximizes voters satisfaction score $\sum_{v_i \in V} \frac{|W \cap C_i|}{|C_i|}$. In the attribute-approval system, we define satisfaction score of v_i for W as $SAV(W, v_i) = \frac{AV(W, \{v_i\})}{min(AV(C, v_i), kd)} = \frac{\sum_{j=1}^{d} |W^j \cap C_i^j|}{min(\sum_{j=1}^{d} |C^j \cap C_i^j|, kd)}$, where W^j (C^j, respectively) is the set of attributes of W (C, respectively) on dimension j. SAV selects the W that maximizes $\sum_{v_i \in V} SAV(W, v_i)$. The complexity of SAV is the same as that of Approval Voting.
Reweighted Approval Voting (RAV) – At every stage, RAV reweighs voter's approval score of a candidate and selects the candidate with the highest approval score [7]. We define reweighed score $RAV(c, v_i)$ as $r(v_i) \times \frac{\sum_{j=1}^{d} |\{v_i | c[j] \in C_i^j\}|}{d}$ where $r(v_i) = \frac{1}{1 + AV(W, v_i)}$. We start with $W = \emptyset$ and at every stage we select a candidate c that maximizes $\sum_{v_i \in V} RAV(c, v_i)$ till $|W| = k$. RAV is a multi-stage AV, and hence, the score computation needs to be done k times. Therefore, the overall computation required in RAV is $O(k(n_a n + md))$.
Proportional Approval Voting (PAV) – The objective of PAV [20] is to maximize the sum of voters' utilities, where the utility of voter v_i is $1 + \frac{1}{2} + \ldots + \frac{1}{|W \cap C_i|}$[2]. With attribute-approval, PAV selects $W \subseteq C$ of size k that maximizes $\sum_{v_i \in V} u(AV(W, v_i))$ where $u(p) = 1 + \frac{1}{2} + \ldots + \frac{1}{\lfloor p \rfloor} + \frac{1}{p}(p - \lfloor p \rfloor)$. PAV is known to be NP-hard [4, 33].
Minimax Approval Voting (MAV)– MAV [8] selects a committee W that minimizes the maximum Hamming distance between W and voters' approval ballots. We define MAV-score of a committee W as $Max(f(W, v_1), f(W, v_2), \ldots, f(W, v_n))$ where $f(W, v_i) = Max(\mathbf{d}(W^j, C_i^j)_{j=1}^{d})$ and $\mathbf{d}(A^j, B^j) = \frac{|A^j \setminus B^j| + |B^j \setminus A^j|}{|D^j|}$. MAV returns a committee W with the lowest MAV-score. MAV is also known to be NP-hard problem [27].

The rules defined in this section and the subsequent sections assume the equal weightage to the various attributes. However, these rules are easily extendable for variable attribute weights set by a centralized authority or expressed by the voters themselves. For instance, a voter v_i may have a weightage $w_i[j]$ for dimension j. In that case, we define the approval voting rule as follows. $Weighted\text{-}AV(c, V) = \sum_{v_i \in V} \sum_{j=1}^{d} (w_i[j] \mid c[j] \in C_i^j)$. Similarly, the other rules also can be extended trivially to a variable weightage setting.

4.2 Properties

In this subsection, we review some standard properties that are desired to be satisfied by multi-winner approval based rules [16]. The summary of different rules and their properties is given in Table 3, in Section 7.
Homogeneity – A rule is said to satisfy the *homogeneity* property if it selects the same W independent of number of times voters' ballot $\mathcal{B} = \{\{C_i^j\}, \forall j \in [1, d], v_i \in V\}$ is replicated.
Consistency – A rule is *consistent* if it satisfies the following implication. If the winning committee is the same W w.r.t. voter lists V and V' individually then it should be the same W with respect to the voter list $V \cup V'$.

[2] Zero if $|W \cap C_i|$ is zero.

Monotonicity– A rule is *monotonic* if it satisfies the following two conditions, 1) If $c \in W$ with respect to V then $c \in W$ with respect to $V_{c[j]} \leftarrow V_{c[j]} \cup V_x, V_x \nsubseteq V_{c[j]}, j \in [1, d]$ 2) If $c \notin W$ with respect to V then $c \notin W$ with respect to $V_{c[j]} \leftarrow V_{c[j]} \setminus V_x, V_x \subseteq V_{c[j]}, j \in [1, d]$.
Committee Monotonicity – Suppose W and W' are the committees selected by rule R with $|W| = k$ and $|W'| = k + 1$. The rule R is *committee monotonic* if $W \subset W'$.

4.3 Unanimity

In vote based committee selection, unanimity refers to an agreement by all voters. Specifically, satisfying the property means that, if there exists a set of candidates who are unanimously approved by all voters then at least one of them should be present in the selected committee W. A rule satisfies unanimity if it selects a committee W such that $\underset{v_i \in V}{\cap} C_i \cap W \neq \emptyset$ when $\underset{v_i \in V}{\cap} C_i \neq \emptyset$ [3].

Using the above definition, we give two definitions of unanimity for attribute level committee selection as follows: 1) *Weak Unanimity* – If $\exists j \in [1, d]$ with $\underset{v_i \in V}{\cap} C_i^j \neq \emptyset$ then $\exists j' \in [1, d]$ with $\underset{v_i \in V}{\cap} C_i^{j'} \cap W^{j'} \neq \emptyset$ where W^j is set of attributes of W on dimension j. 2) *Strong Unanimity* – $\forall j \in [1, d]$ with $\underset{v_i \in V}{\cap} C_i^j \neq \emptyset$ it holds $\underset{v_i \in V}{\cap} C_i^j \cap W^j \neq \emptyset$. We note that for $d = 1$ both weak and strong unanimity[3] convey the same meaning.

Lemma 4.1. *For $k < d$, there may not exist a committee that satisfies strong unanimity.*

Proof. Consider two candidates $c_1 = [a_1, b_1]$ and $c_2 = [a_2, b_2]$, and $k = 1$. The approvals for each of these attributes are given as $V_{a1} = V_{b2} = V$ and $V_{a2} = V_{b1} = \{v_1\}$. Selecting any one of these two candidates violates the strong unanimity property. But, one can assure a committee W that provides strong unanimity when $k \geq d$. □

Lemma 4.2. *AV, SAV, RAV, PAV, and MAV do not satisfy weak unanimity for $k \geq 1$ and $d > 1$ and they satisfy weak unanimity for $d = 1$.*

Proof. Let $X^1 = V \setminus \{v_n\}$ and $X^2 = V \setminus \{v_1\}$. $V_{c_1[1]} = V$, $(V_{c_1[j]} = v_1)_{j=2}^d$, $((V_{c_i[j]} = X^1)_{i=2}^{\lceil m/2 \rceil})_{j=1}^d$ and $((V_{c_i[j]} = X^2)_{i=\lceil m/2 \rceil+1}^m)_{j=1}^d$. AV, or SAV, or RAV, or PAV, or MAV selects a set $W \subseteq C \setminus \{c_1\}$, whereas $c_1[1]$ is the only attribute, which is unanimously approved by all voters and is not part of W. Hence, AV, SAV, RAV, PAV, and MAV do not satisfy weak unanimity for $d > 1$. When $d = 1$, the analysis is the same as given in [3]. □

Proposition 4.3. *If a rule does not satisfy weak unanimity then it does not satisfy strong unanimity as well. Hence, none of the extended rules satisfy unanimity.*

Besides the above properties, *Justified Representation* is a desirable property of approval voting based rules. We study and present a detailed analysis of this property in the next section.

[3]Going further, if there exist multiple unanimous candidates, then all of them should be present in the committee W.

5 JUSTIFIED REPRESENTATION

Justified Representation (JR)[4] is a desirable property of approval based rules in many real-world committee selections. For satisfying this representation, if there exists a sizeable group of voters with common preferences then the group should have representation in the committee. Based on this, we provide two definitions of justified representation for attribute-approval voting as follows: 1) Simple Justified Representation (SJR) and 2) Compound Justified Representation (CJR). For $d = 1$, simple and compound justified representations are the same.

Definition 5.1. A committee W satisfies simple justified representation if $\forall V' \subseteq V$: $(|V'| \geq \frac{n}{k}) \wedge (\exists j \in [1, d]\ with \underset{v_i \in V'}{\cap} C_i^j \neq \emptyset) \implies (\exists j' \in [1, d]\ with \underset{v_i \in V'}{\cup} C_i^{j'} \cap W^{j'} \neq \emptyset)$.

For $k = 1$, any random candidate also satisfies simple justified representation unless there exists a candidate who has no approval on any attribute.

Lemma 5.2. *Approval voting does not satisfy SJR for $k \geq 2$ and $d \geq 2$ or when $k \geq 3$ and $d \geq 1$.*

Proof. Let $C = \{c_1, c_2, \dots, c_{k+2}\}$ be the set of candidates, $X^1 = \{v_i, i \in [1, n-\frac{n}{k}]\}$, and $X^2 = \{v_i, i \in [n-\frac{n}{k}+1, n]\}$. Consider profiles $((V_{c_i[j]} = X^1)_{i=1}^k)_{j=1}^d$, $(V_{c_{k+1}[j]} = \{X^2\})_{j=1}^{\frac{d}{2}}$, $(V_{c_{k+1}[j]} = v_n)_{\frac{d}{2}+1}^d$, $(V_{c_{k+2}[j]} = v_n)_{j=1}^{\frac{d}{2}}$ and $(V_{c_{k+2}[j]} = \{X^2\})_{j=\frac{d}{2}+1}^d$. Candidate c_{k+1} or c_{k+2} has to be present in W in order to satisfy SJR whereas AV selects $\{c_1, \dots, c_k\}$. It is proven in [3] that AV fails JR for $k \geq 3$, and when $d = 1$, SJR is the same as JR. Hence, AV fails SJR for $k \geq 3$ and $d \geq 1$. □

Lemma 5.3. *PAV and RAV do not satisfy SJR for $k \geq 3$ and $d \geq 2$ or $d \geq 3$ and $k \geq 2$. For $k \leq 2$ and $d \leq 2$, PAV and RAV satisfy SJR.*

Proof. We omit the generalized proof due to its complexity. Let $C = \{c_1, c_2, \dots, c_5\}$, $n = 90$, $d = 2$, and $k = 3$, attribute-wise voter lists are $(V_{c_1[j]} = \{v_i\}_{i=1}^{35})_{j=1}^d$, $(V_{c_2[j]} = \{v_i\}_{i=21}^{55})_{j=1}^d$, $(V_{c_3[j]} = \{v_i\}_{i=26}^{50})_{j=1}^d$, $(V_{c_4[j]} = \{v_1\})_{j=1}^{\frac{d}{2}}$, $(V_{c_4[j]} = \{v_i\}_{i=61}^{90})_{j=\frac{d}{2}+1}^d$, $(V_{c_5[j]} = \{v_i\}_{i=61}^{90})_{j=1}^{\frac{d}{2}}$, $(V_{c_5[j]} = \{v_1\})_{j=\frac{d}{2}+1}^d$. PAV (or, RAV) selects $\{c_1, c_2, c_3\}$ and ignores set of $\frac{n}{k}$ voters who jointly approved $c_4[2]$ and $c_5[1]$. To extend the proof to $k > 3$, we take $k - 3$ additional candidates and $(k - 3) \times 30$ additional voters, and assign 30 unique votes to each attribute of a new candidate. For $k = 2$ and $d \geq 3$: Let $(V_{c_1[j]} = V_{c_2[j]} = \{v_i\}_{i=1}^{30})_{j=1}^d$, $V_{c_3[1]} = \{v_i\}_{31}^{60}$, $(V_{c_3[j]} = \{v_1\})_{j=2}^d$. PAV or RAV selects W as $\{c_1, c_2\}$ and disregards a set of n/k voters. Hence, PAV and RAV do not satisfy SJR for $k = 2$ and $d \geq 3$. Similarly for $k > 2$ and $d \geq 3$. For $k = 2$ and $d = 2$, RAV selects a candidate with highest AV score in the first iteration. The candidate having $\frac{n}{2}$ approvals for one of its attributes will have the highest score in the second iteration. Similar logic works for PAV. Hence, RAV and PAV satisfy SJR for $k = 2$ and $d = 2$. When $d = 1$, the analysis is the same as given in [3]. □

[4]Similar notions are, Extended JR [3], Proportional JR [31], Proportional Representation [30] and Strong Proportional Representation [30].

LEMMA 5.4. *SAV and MAV do not satisfy SJR for $k \geq 2$, $d \geq 1$.*

Proof. When $d = 1$, attribute-approvals can be visualized as candidate-approvals. In the case of candidate-approvals, MAV and SAV do not satisfy justified representation for $k \geq 2$ [3]. If we replicate the voters' ballots across all the dimensions then the proof follows. □

Definition 5.5. W provides CJR if $\forall\ V' \subseteq V$ and $\forall i \in [1, d]$: $(|V'| \geq \frac{n}{k}) \wedge (\underset{v_i \in V'}{\cap} C_i^j \neq \emptyset) \implies (\underset{v_i \in V'}{\cup} C_i^j \cap W^j \neq \emptyset)$.

LEMMA 5.6. *For $k \geq 3$, there may not exist a committee that provides compound justified representation .*

Proof. Consider a set of candidates $c_i = [a_i, b_i], i \in [1, m]$ with $m = 6$. Voting approvals of attributes are given as $V_{a1} = V_{b4} = \{v_1, v_2\}$, $V_{a2} = V_{b5} = \{v_3, v_4\}$, $V_{a3} = V_{b6} = \{v_5, v_6\}, V_{b1} = V_{b2} = V_{b3} = V_{a4} = V_{a5} = V_{a6} = \{v_1\}$. For $k = 3$, none of the three candidate committees satisfy compound justified representation. One can assure a committee that satisfies CJR for $k = 2$, if we assume that each attribute is approved by at least one voter. □

PROPOSITION 5.7. *CJR implies SJR.*

LEMMA 5.8. *Checking whether there exists a committee that provides CJR is a NP-complete problem for $k \geq 2$ and $d \geq 2$.*

Proof. Given a committee, CJR satisfiability is verifiable in polynomial time. Hence, the problem is in NP. We reduce the set cover problem to the committee selection problem to show that the current problem is NP-hard. Given a set of subsets $S_1, S_2, \ldots, S_{m'}$ with n' elements, the set-cover problem is to select k' subsets such that they cover all the elements in $\cup_{i=1}^{m'} S_i$. We take an instance of a set cover problem with n' elements and m' subsets, and construct $m = m' + 2kn'$ candidates with $n = kn'$ voters. We take elements of a subset sum problem as voters and subsets of elements as subsets of voters. Let $V_1^{CJR} = \cup_{i=1}^{m'} S_i = \{v_1, v_2, \ldots, v_{n'}\}$ be the subset of voters constructed from the elements of a subset selection problem and let $V_i^{CJR} = \{v_{n'(i-1)+1}, \ldots, v_{n'i}\}, \forall i \in [2, k]$. We construct candidates' approvals as follows using the voter list $V = \cup_{i=1}^{k} V_i^{CJR}$.
$Initialize: V_{c_i[j]} \leftarrow \emptyset, \forall\ i \in [1, m], j \in [1, 2]$;
$V_{c_i[j]} \leftarrow S_i, \forall\ i \in [1, m'], j \in [1, 2]$; $V_{c_x[1]} \leftarrow \{v_i\} \cup \{V_h^{CJR} \setminus v_{hn'}\}, \forall\ i \in [1, n'], h \in [2, k]$, where $x = m' + (h-2)n' + i$. $V_{c_x[2]} \leftarrow \{v_i\} \cup \{V_h^{CJR} \setminus v_{hn'}\}, \forall\ i \in [1, n'], h \in [2, k]$, where $x = m' + (k+h-3)n' + i$. $V_{c_x[1]} \leftarrow \{V_1^{CJR} \setminus v_i\} \cup \{v_{2n'}\}, \forall\ i \in [1, n']$, where $x = m' + (2k-2)n' + i$. $V_{c_x[2]} \leftarrow \{V_1^{CJR} \setminus v_i\} \cup \{v_{2n'}\}, \forall\ i \in [1, n']$, where $x = m' + (2k-1)n' + i$. $\forall\ i \in [m+1, m+2kn'], j \in [1, 2]$ *and* $k \geq 3$, *if* $V_{c_i[j]} = \emptyset$ *then* $V_{c_i[j]} = \{v_{hn'}\}, h \in [3, k]$. We can see that if there is a yes instance in the set cover problem, there will be a yes instance in our problem for $k \geq 2$, otherwise not. One can try the following instance of a subset-sum problem to work out the reduction. $S_1 = \{v_1, \ldots v_7\}$, $S_2 = \{v_8, \ldots v_{14}\}$, $S_3 = \{v_1, \ldots v_4, v_8 \ldots v_{11}\}$, $S_4 = \{v_5, v_6, v_{12}, v_{13}\}$, and $S_5 = \{v_7, v_{14}\}$ are subsets from the set cover problem with $n' = 14$, $m' = 5$, and $k = 2$. □

We adapt Greedy Approval Voting (GAV) [3] and extend it to attribute level in the next section. We show that GAV satisfies unanimity, SJR, and CJR under some assumptions.

5.1 Justified Committee with highest (S)AV

We note that the committee with the highest (S)AV may not always be a justified committee (as shown in Lemmas 5.2 and 5.4). Lemma 5.9 illustrates that many justified committees are possible for the given voter approvals, we take an example of $d = 1$ for the simplicity of perception. Further, we prove the following two problems to be NP-complete, 1) Selection of justified committee with highest AV and 2) Selection of justified committee with highest SAV. The preceding discussion is applicable for both simple and compound justified representations. However, we have already proven that finding a compound justified representation committee problem is NP-complete for $k \geq 2$ and $d \geq 2$, and when d=1, compound and simple justified representations are the same. Therefore, we give the NP-completeness proof related to simple justified representation in the Lemma 5.10.

LEMMA 5.9. *More than one justified committee may exist.*

PROOF. Let us consider a simple example with 10 voters, 6 candidates and $k = 2$. Voters' approvals for each candidate is given as $V_1 = \{v_1, v_2, v_3, v_4, v_5\}, V_2 = \{v_6, v_7, v_8, v_9, v_{10}\}, V_3 = \{v_1, v_3, v_5\}, V_4 = \{v_6, v_7, v_9\}, V_5 = \{v_2, v_4, v_6, v_8, v_{10}\}, V_6 = \{v_2, v_6\}$. In this example, any two of six candidates forms a justified committee. □

5.1.1 *Justified committee with highest Approval Voting.*

Approval Voting (AV) rule satisfies *Justified Representation* for $k = 1 \wedge d = 1$ and $k = 2 \wedge d = 1$, if ties are broken in favor of sets that provide justified representation. However, it is not the case for $k \geq 2 \wedge d \geq 2$ and $k \geq 3 \wedge d = 1$.

LEMMA 5.10. *Justified committee with highest AV problem is NP-complete for $k \geq 2 \wedge d \geq 2$ and $k \geq 3 \wedge d = 1$.*

PROOF. Let V_W be a multi-set representing all the voters who approved attributes of a committee W. If a voter approves multiple attributes in W he would appear multiple times in V_W. The following is the decision problem corresponding to the justified committee with highest AV problem.
Question: Does there exist a committee W of size k such that $|V_W| \geq \tau$ and $|V_{c_i[j]} \setminus V_W| < \frac{n}{k}$, $\forall c_i[j] \in \{C \setminus W\}$?.

Given a certificate of solution, we can easily verify that whether or not a committee satisfies the required constraints in polynomial time. Hence, the problem is in NP. We reduce the set cover problem to the committee selection problem to show that the current problem is NP-hard. To reduce set cover problem, we take an instance of a set cover problem with n' number of elements and m' subsets. For every instance with n' elements and m' subsets in a subset cover problem, we have $n = 2kn'$ number of voters and $m = m' + n'$ number of candidates in a committee selection problem. Consequently, we have *md* subsets in the committee selection problem. The subsets in the committee selection problem are the set of voters who approved for each attribute-value. Let $S_i, i \in [1, m']$ are subsets from the set cover problem, V' and V'' are disjoint voter lists with $|V'| = (2k-3)n'$ and $|V''| = 2n'$ respectively. Let $S = \cup_{i=1}^{m'} S_i = \{v_1, v_2, \ldots, v_{n'}\}$. We construct the candidates' subsets/voter-lists in the following manner. We take m' subsets with n' elements from the set cover problem and add a set V' of voters to each subset i.e., $V_{c_i[j]} \leftarrow \{S_i \cup V'\}, \forall i \in [1, m'], j \in [1, d]$. Next, we construct n' additional candidate voter-lists with one voter from S in each of these lists, i.e., $V_{c_i[1]} \leftarrow \{\{v_{i-m'}\} \cup \{V''_{2n'-1}\}\}$ *and* $V_{c_i[j]} \leftarrow \{v_{i-m'}\}, \forall i \in [m'+1, m'+n'], j \in [2, d]$, where ${V''}_{2n'-1}$ is a proper subset of V'' with size $2n'-1$. Finally, we set τ to $d(k((2k-3)n')+n')$. It is easy to see that there exist a justified committee of size k with approval count greater than or equal to τ iff there exists a set cover, for $k \geq 2$ and $d \geq 2$. The same proof works for $k \geq 3$ and $d = 1$ if we exclude the rules relevant for $d > 1$. However, NP-complete proof related to $k \geq 3 \wedge d = 1$ is invented simultaneously [10]. □

To illustrate, let us take the following instance of a set cover problem with $n' = 14$, $m' = 5$, $S_1 = \{v_1, \dots v_7\}$, $S_2 = \{v_8, \dots v_{14}\}$, $S_3 = \{v_1, \dots v_4, v_8 \dots v_{11}\}$, $S_4 = \{v_5, v_6, v_{12}, v_{13}\}$, $S_5 = \{v_7, v_{14}\}$. We create $m = m' + n' = 19$ number of candidates in a committee selection problem with a set of voters approved for each candidate, and we set $d = 2$ and $k = 2$. Let $V' = \{v_{15}, \dots, v_{28}\}$ and $V'' = \{v_{29}, \dots, v_{56}\}$. We create voter sets, $V_{c_i[j]}, i \in [1, 19], j \in [1, 2]$, corresponding to all m candidates over d dimensions using the above rules. We then set $\tau = d(k((2k-3)n')+n') = 2(2(14)+14) = 84$, and $\frac{n}{k} = 2n' = 28$. If there exists a justified committee of size $k = 2$ with approval score ≥ 84 then there is a set cover of size k in the set cover. Otherwise, there is no set cover of size k and vice-versa.

5.1.2 ***Approximation of highest AV.*** We propose a heuristic to find a justified committee with highest AV and analyze different aspects of the proposed heuristic. Algorithm 1 describes the procedure. In the algorithm, V_* denotes a set of voters who approved at least one attribute of a candidate c_*. The *Descending-Sorted-list-of-candidates* method in the algorithm sorts candidates in descending order of their approval scores. The basic idea of our heuristic is to select the candidate with highest approval score at every level such that the algorithm checks for maximum allowable unrepresentative voters to make a set closer to a justified representation.

Algorithm 1: Justified Committee with highest AV

Input: $C, V, k, \{C_i^j, \forall i, j\}$ **Output:** $W \subseteq C, |W| = k$
$W \leftarrow \emptyset$; $V' \leftarrow V$; Compute $AV(c_i, V)$ for each $c_i \in C$;
$C' \leftarrow$ Descending-Sorted-list-of-candidates($AV(c_i, V), \forall c_i$);
for *l = 1 to k* **do**
  $i \leftarrow 1$;
  do $c_* \leftarrow C'[i]$; $V'' \leftarrow V' \setminus V*$; $i \leftarrow i+1$;
  while $\left|\left\{ \bigcup_{\forall i,j} V''_{c_i[j]} : |V''_{c_i[j]}| \geq \frac{n}{k} \right\}\right| \geq \frac{n(k-l+1)}{k}$
  $W \leftarrow W \cup \{c_*\}$; $C' \leftarrow C' \setminus \{c_*\}$; $V' \leftarrow V' \setminus V*$;
end

Note that, the proposed heuristic always yields higher or equal approval voting score set than GAV [3].

Lemma 5.11. *Let S^* be the highest score that could be obtained by satisfying the justified representation property and S be the score given by our approximation scheme. $\frac{S^*}{S} \leq 2, \forall k \geq 1, d = 1$.*

Analysis: The ratio between the ground truth and approximate score is maximum when the highest-scoring candidates cover a sufficiently more sizable crowd to satisfy justified representation. However, the greedy way of our selection failed to ensure that justified representation is satisfied with the subset of candidates. Each time when the algorithm fails, it selects a candidate with at least n/k number of votes. The more times the algorithm fails, the ratio would be higher. The highest possible approval count for an individual candidate if the algorithm fails at step $l+1$ is $dl(\frac{n}{k}-1)$. If the algorithm fails at step $l+1$ first time, the maximum number of times that the algorithm would fail is $(k-l)$ and $2 \leq l \leq k$. Therefore, the task is to find a value of l for which the ratio is maximum. We formulate this problem as follows.

$$
\begin{aligned}
\frac{S^*}{S} &= \max_{2 \leq l < k} \frac{dl(\frac{n}{k}-1)k}{l(\frac{n}{k}-1)(l) + (\frac{n}{k})(k-l)} \\
&< \max_{2 \leq l < k} \frac{dl(\frac{n}{k})k}{dl(\frac{n}{k})(l) + (\frac{n}{k})(k-l)}, \forall\, k > 1 \\
&= \max_{2 \leq l < k} \frac{dlnk}{dlnl + n(k-l)}
\end{aligned}
$$

The above quantity varies based on k and d values. We note that, when $k = 2$ and $d = 1$ the fraction $\frac{S^*}{S} = 1$, which means that our approach is optimal for $k = 2$ and $d = 1$. For $k = 3$ and $d = 1$, it is 1.2, for $k = 4$ and $d = 1$, it is 1.33, and so forth. The best case occurs when $k = l$, $\frac{S^*}{S} = 1$. Similarly, worst case occurs when $l = 2$ and average case when $l = (k+2)/2$.

Complexity Analysis: Construction of a voter list for each attribute-value from an attribute approvals voters and the computation of AV score is $O(n_a n + md)$ complexity. Sorting m candidates according to the size of their voter-list is $O(mlogm)$. The complexity of for loop is $O(mkd)$. Hence, the overall complexity is $O(n_a n + mlogm + mkd)$.

5.1.3 ***Justified Committee with highest SAV.*** Satisfaction Approval Voting (SAV) yields the same results as Approval Voting when all voters approve an equal number of candidates. Hence, as a direct extension, we note that finding the justified committee with the highest-SAV problem is NP-complete.

5.2 Proportional Justified Representation

The motivation of Proportional Justified Representation (PJR) is that the larger groups with more joint approvals should have more representatives in the committee W. The group's size determines the number of representatives that a group should have in the committee W. We extend this concept to attribute level preferences and define the following two properties.

Definition 5.12. W provides simple proportional justified representation if $\forall V' \subseteq V : (|V'| \geq \frac{ln}{k}) \wedge (\exists j \in [1, d]\ s.t. \bigcap_{v_i \in V'} C_i^j \geq l) \implies (\exists j' \in [1, d]\ s.t. |\bigcup_{v_i \in V'} C_i^{j'} \cap W^{j'}| \geq l), \forall l \in [1, k]$.

Definition 5.13. W provides compound proportional justified representation if $\forall V' \subseteq V$ and $\forall j \in [i, d] : (|V'| \geq \frac{ln}{k}) \wedge (\bigcap_{v_i \in V'} C_i^j \geq l) \implies (|\bigcup_{v_i \in V'} C_i^j \cap W^j| \geq l), \forall l \in [1, k]$.

When $l = 1$, simple proportional justified representation (compound proportional justified representation, respectively) is the same as simple justified representation (compound justified representation, respectively). Hence, from Lemma 5.6, we can say that a committee that provides compound proportional justified representation may not exist. Also, checking whether there exist a committee that provides Compound PJR (CPJR) is NP-complete for $k \geq 2$ and $d \geq 2$. The proof follows directly from Lemma 5.8.

Proposition 5.14. *1) If a rule does not satisfy a simple justified representation, it will not satisfy simple proportional justified representation as well as compound justified representation. 2) Similarly, if a rule does not satisfy a compound justified representation, it will not satisfy compound proportional justified representation. 3) Moreover, If a rule does not satisfy simple(or compound) justified representation for $d = 1$, it would not satisfy the same for $d \geq 1$. Hence, from 1, 2, and 3, we can derive the following conclusions. SAV and MAV do not provide simple PJR and compound PJR for $k \geq 2$. AV does not satisfy simple PJR for $k \geq 2$ and $d \geq 2$. Finally, PAV and RAV do not satisfy simple PJR for $k \geq 3$ and $d \geq 2$ or $k \geq 2$ and $d \geq 3$. It is easy to see that they satisfy simple PJR for $k \leq 2$ and $d \leq 2$.*

6 GREEDY APPROVAL VOTING

Attribute level *Greedy Approval Voting (GAV)*, shown in Algorithm 2, starts by setting $V' = V$ and $W = \emptyset$. At each iteration, GAV selects a candidate c_* having an attribute with the highest number of approvals with respect to V' and add it to W. GAV removes all voters who voted for at least one attribute of c_* from V'. This process is repeated till $|W| = k$. In case the voter list V' is empty when $|W| < k$, we set V' to V. Once the voter list is empty, random selection of candidates would satisfy the weak unanimity and SJR properties but fails many other properties.

Algorithm 2: Greedy Approval Voting

Input: $C, V, k, \{C_i^j, \forall v_i \in V, j \in [1, d]\}$
Output: $W \subseteq C,\ |W| = k$
$W \leftarrow \emptyset; V' \leftarrow V;$
while $|W| < k$ **do**
  $c_* \leftarrow \underset{c_i \in C}{Argmax}\ Max(|V'_{c_i[j]}|)_{j=1}^{d};$
  $W \leftarrow W \cup \{c_*\}; C \leftarrow C \setminus \{c_*\}; V' \leftarrow V' \setminus \{\cup_j V_{c_*[j]}\};$
  if $|V'| = 0$ **then** $V' \leftarrow V;$
end

Lemma 6.1. *Greedy Approval Voting satisfies weak unanimity.*

Proof. If there exists an attribute which is unanimously approved by all voters, GAV selects the corresponding candidate first. Hence, GAV satisfies weak unanimity. □

Lemma 6.2. *GAV satisfies strong unanimity if ties are broken in favor of the candidates that provide strong unanimity.*

Proof. We say that a voter is unrepresented on dimension j if none of his/her approved attributes from the domain D^j are present in W. We define the following tie-breaking rules: 1) If there are multiple attributes with the same number of approvals, we select the one with the highest number of approvals according to the unrepresentative voters of the dimension where the attribute is present. 2) If multiple candidates have a unanimous attribute, we select the one with the highest number of unanimous attributes. Using these two rules, GAV selects a committee with at least one unanimous attribute on every dimension (if there is such an attribute on that dimension). Hence, GAV satisfies strong unanimity. □

Lemma 6.3. *GAV satisfies simple justified representation.*

Proof. A rule does not satisfy simple justified representation if it completely ignores a set of $\frac{n}{k}$ voters who jointly approved for some attribute. If we prove that GAV does not leave any $\frac{n}{k}$ voters who jointly approved for an attribute, we can say that GAV satisfies simple justified representation. GAV is a multi-stage approach, and at every step, it selects a candidate having an attribute with a maximum number of approvals from unrepresented voters. Even if there are entirely disjoint sets of voters, each size $\frac{n}{k}$, GAV can cover all such voters in k steps. Hence, GAV satisfies simple justified representation. □

In Section 5, we have shown that CJR is an NP-complete problem. However, one can solve it in polynomial time under certain assumptions. We consider each dimension separately and identify a set of attributes that satisfy justified representation for that dimension. This can be done in polynomial time using the proposed GAV. Let J^i be the set of attributes that satisfies justified representation for dimension i. Selection of a committee that satisfies CJR is polynomial if we assume that $J^1 \times J^2 \times \ldots \times J^d \subseteq C$. In addition to unanimity and justified representation, GAV satisfies other properties described in Section 3. Nevertheless, GAV does not satisfy simple proportional justified representation and compound proportional justified representation.

7 JUSTIFIED APPROVAL VOTING

We propose a new polynomial-time rule called Justified Approval Voting (JAV) that satisfies Simple Proportional Justified Representation (SPJR). First we discuss the algorithm for $d = 1$, i.e., for a committee selection problem with candidate approvals. We then extend the algorithm to $d > 1$, i.e., to attribute approvals. Algorithm 3 gives a detailed procedure of JAV for $d = 1$.

Algorithm 3: Justified Approval Voting

Input: $C, V, k, \{C_i, \forall v_i \in V\}$
Output: $W \subseteq C, |W| = k$
$W \leftarrow \emptyset;$
for *l = 1 to k* **do**
  $V' \leftarrow V \setminus \{v_i | v_i \in V \wedge |C_i \cap W| \geq l\};$
  $c_* \leftarrow \underset{c_i \in C}{ARGMAX}\ AV(c_i, V');$
  $W \leftarrow W \cup \{c_*\}; C \leftarrow C \setminus \{c_*\};$
  while $|W| < k \wedge AV(c_*, V') \geq \frac{ln}{k}$ **do**
    $V' \leftarrow V' \setminus V_*; c_* \leftarrow \underset{c_i \in C}{ARGMAX}\ AV(c_i, V');$
    if $AV(c_*, V') < \frac{ln}{k}$ **then break;**
    $W \leftarrow W \cup \{c_*\}; C \leftarrow C \setminus \{c_*\};$
  end
  if $|W| == k$ **then exit;**
end

We use the following definitions to prove the correctness of the proposed JAV. *l-Representative Voter* – A voter is said to be l-representative if at least l of his/her approved candidates are present in the selected committee W. *l-Unrepresentative Voter* – A voter is said to be l-unrepresentative if less than l of his/her approved candidates are present in the selected committee W.

Lemma 7.1. *Justified Approval Voting satisfies simple proportional justified representation .*

Proof. To prove the algorithm's correctness, we need to show that every $\frac{ln}{k}$ group of voters set having l joint approvals have no less than l representatives in the committee W, $1 \leq l \leq k$. If $\frac{ln}{k}$ voters have at least l joint approvals and none of these voters are l-representative at level l, there is at least one candidate with $\geq \frac{ln}{k}$ approvals. In every step ($l = 1$ to k), JAV ensures that no candidate with $\geq \frac{ln}{k}$ approvals of l-unrepresentative voters are left out from selection. Hence, JAV ensures that at least l representatives from every group of $\frac{ln}{k}$ voters who jointly approved l candidates. □

We extend the JAV algorithm described above to the attribute approvals by iteratively using the Algorithm 3 on the individual attributes with domain values of an attribute as candidates. Algorithm 4 outlines the procedure. The only additional constraint imposed here is that at level l, no candidate will be selected with $< \frac{ln}{k}$ approvals. If there are no such candidates, the algorithm terminates without necessitating the selection of k candidates. These attribute level committees are then sorted in the descending order according to their size. For each of these committees, starting from the largest one, the *Select_Candidates* subroutine looks for l attribute-values having approvals $\geq \frac{ln}{k}$ for some value of l in the reverse range of $[1, k]$. It then selects the candidates corresponding to these l-attribute values[5]. After that, we identify *l-representative voters* and remove them from the voter list. We repeat the candidate selection with the new voter list and by decrementing the value of l until $|W| = k$. If there are no such k candidates, the algorithm may select candidates based on a different strategy, even the random candidates serves the purpose.

Algorithm 4: Attribute Level Justified Approval Voting

Input: $C, V, k, \{C_i^j, \forall i, j\}$ **Output:** $W \subseteq C, |W| = k$ **for** j = *1 to* d **do** $W\{j\} \leftarrow JAV(D^j, V, k, \{C_i^j, \forall v_i, j\})$;
$SI \leftarrow Sorted\_Indices(W)$;
$W \leftarrow Select\_Candidates(W\{j\}_{j=1}^{d}, SI, C, l \in [k, 1])$;

Sorting the individual committees and selecting l in the reverse order ensures the representation of candidates with higher joint approvals first. Here, the representation of higher joint approvals ahead is significant to cover nested joint approvals. Removal of *l-representative* voters in every iteration guarantees the selection of another smaller group that have disjoint approvals with the previous group. Also, assures that every $\frac{ln}{k}$ voters with l joint approvals have l representatives on at least one dimension. Therefore, JAV satisfies SPJR, consequently SJR. Besides, JAV satisfies committee selection's other desirable properties as shown in Table 3. In the table, '$*$' indicates that a rule satisfies the corresponding property P if there is a committee that satisfies both P and SPJR, as well as ties, are broken in favor of that committee.

Complexity Analysis: We first analyze the complexity of Algorithm 3 with respect to candidate approvals and then we analyze it for attribute level approvals (Algorithm 4). Score computation of each candidate involves scanning n voters' approval ballots and we compute the score for m candidates. Hence, complexity for score computation is $O(mn)$. Score computation is taking place at two places in the algorithm (line 4 and line 10). It is to be noted that if the second loop repeats for k number of times, the first loop would repeat single time only and vice versa. On the whole, score computation happens only k number of times. Therefore the complexity of the algorithm is $O(knm)$. Best case occurs when there is no change in the voter list at every iteration. In that case, implementation can take care of avoiding unnecessary score computation. This may happen in two situations, 1) Every item is approved by less than n/k voters and 2) Majority of the voters approves all the k candidates. Hence, The complexity of the best case is $O(nm)$, and in average and worst cases it is $O(knm)$. Furthermore, the complexity of attribute level JAV is $O(d \times (nm_D))$ in the best case and $O(d \times (knm_D))$ in the worst and average cases, where m_D is the maximum domain size among attributes. We note that $m_D = m$ when $d = 1$.

[5]If there is more than one candidate with the same attribute value, we prefer a candidate that covers the most representative attributes from other committees.

Table 3: Summary of the properties satisfied by rules.

	AV	SAV	PAV	RAV	MAV	GAV	JAV
Homogeneity	✓	✓	✓	✓	✓	✓	✓
Consistency	✓	✓	✓	✓	✓	✓	✓
Monotonicity	✓	✓	✓	✓	✓	✓	✓
Committee Monotonicity	✓	✓	×	✓	×	✓	✓
Weak Unanimity	×	×	×	×	×	✓	✓*
Strong Unanimity	×	×	×	×	×	✓	✓*
SJR	×	×	×	×	×	✓	✓
SPJR	×	×	×	×	×	×	✓
CJR or CPJR	×	×	×	×	×	×	×

8 CONCLUSIONS AND DISCUSSION

The present work initiates a new direction of research, namely, attribute approvals for a committee selection problem. The current proposal is an initiation for a committee selection problem with attribute level preferences, and there is much scope for future research. Extending the problem with constraint satisfaction is one of the potential problems for future research. In this work, we have seen one way of extending the voters' preferences on attributes to candidates or committees, i.e., using scoring rules. Examining different ways of transforming voters' approvals over attributes to the committees is an excellent direction for future research. For instance, we can represent a committee W as a vector $(x_1, \ldots, x_d)$ wherein x_i is the number of approvals that the attribute i has got in committee W. This preference extension induces a partial order for each voter over the candidate set. One can then use this preference relation over committees to reason about axioms such as Pareto-optimality. Finally, we extended the standard representation axioms that reflect how fairly the outcome represents the voters. It will be interesting to simultaneously focus on fair representation of attributes. In many real world applications, there may be different weightings for candidates satisfying different attributes. We leave this important extension for future work.

Acknowledgments: We thank Dr. Bruhadeshwar (IIT Jammu) for his valuable discussions and guidance in improving the writing/presentation aspects of the manuscript.

REFERENCES

[1] Gilbert Abotisem Abiiro, Gerald Leppert, Grace Bongololo Mbera, Paul J Robyn, and Manuela De Allegri. 2014. Developing attributes and attribute-levels for a discrete choice experiment on micro health insurance in rural Malawi. *BMC health services research* 14, 1 (2014), 1.

[2] Haris Aziz, Felix Brandt, Edith Elkind, and Piotr Skowron. 2019. Computational Social Choice: The First Ten Years and Beyond. In *Computing and Software Science - State of the Art and Perspectives*, Bernhard Steffen and Gerhard J. Woeginger (Eds.). Lecture Notes in Computer Science, Vol. 10000. Springer, 48–65. https://doi.org/10.1007/978-3-319-91908-9_4

[3] Haris Aziz, Markus Brill, Vincent Conitzer, Edith Elkind, Rupert Freeman, and Toby Walsh. 2017. Justified representation in approval-based committee voting. *Social Choice and Welfare* 48, 2 (2017), 461–485.

[4] Haris Aziz, Serge Gaspers, Joachim Gudmundsson, Simon Mackenzie, Nicholas Mattei, and Toby Walsh. 2015. Computational aspects of multi-winner approval voting. In *AAMAS*. 107–115.

[5] Steven J Brams. 1990. Constrained approval voting: A voting system to elect a governing board. *Interfaces* 20, 5 (1990), 67–80.

[6] Steven J Brams and Peter C Fishburn. 1978. Approval voting. *American Political Science Review* 72, 03 (1978), 831–847.

[7] Steven J Brams and D Marc Kilgour. 2014. Satisfaction approval voting. In *Voting Power and Procedures*. Springer, 323–346.

[8] Steven J Brams, D Marc Kilgour, and M Remzi Sanver. 2007. A minimax procedure for electing committees. *Public Choice* 132, 3-4 (2007), 401–420.

[9] Robert Bredereck, Piotr Faliszewski, Ayumi Igarashi, Martin Lackner, and Piotr Skowron. 2018. Multiwinner elections with diversity constraints. In *AAAI*. 933–940.

[10] Robert Bredereck, Piotr Faliszewski, Andrzej Kaczmarczyk, and Rolf Niedermeier. 2019. An Experimental View on Committees Providing Justified Representation.. In *IJCAI*. 109–115.

[11] M. Brill, P. Golz, D. Peters, U. Schmidt-Kraepelin, and K. Wilker. 2019. Approval-Based Apportionment. (2019).

[12] John C Butler, James S Dyer, Jianmin Jia, and Kerem Tomak. 2008. Enabling e-transactions with multi-attribute preference models. *European Journal of Operational Research* 186, 2 (2008), 748–765.

[13] Henny Collen and Joris Hoekstra. 2001. Values as determinants of preferences for housing attributes. *Journal of Housing and the Built Environment* 16, 3-4 (2001), 285–306.

[14] Ernesto Dal Bo. 2007. Bribing voters. *American Journal of Political Science* 51, 4 (2007), 789–803.

[15] Antonio Gomes De Menezes and JC Vieira. 2008. Willingness to pay for airline services attributes: evidence from a stated preferences choice game. *European Transport* 39 (2008), 1–13.

[16] Edith Elkind, Piotr Faliszewski, Piotr Skowron, and Arkadii Slinko. 2017. Properties of multiwinner voting rules. *Social Choice and Welfare* 48, 3 (2017), 599–632.

[17] Abdul Hadi, Ahmad Hanis Izani, Mad Nasir Shamsudin, Jinap Selamat, Alias Radam, Abdul Karim, and Muhammad Shahrim. 2013. Consumer's preferences for eggs attributes in Malaysia: evidence from conjoint survey. *International Food Research Journal* 20, 5 (2013), 2865–2872.

[18] Jac C Heckelman. 1998. Bribing voters without verification. *The Social Science Journal* 35, 3 (1998), 435–443.

[19] David A Hensher. 1994. Stated preference analysis of travel choices: the state of practice. *Transportation* 21, 2 (1994), 107–133.

[20] D Marc Kilgour. 2010. Approval balloting for multi-winner elections. In *Handbook on approval voting*. Springer, 105–124.

[21] Alison M Konrad, J Edgar Ritchie Jr, Pamela Lieb, and Elizabeth Corrigall. 2000. Sex differences and similarities in job attribute preferences: a meta-analysis. *Psychological bulletin* 126, 4 (2000), 593.

[22] Marit E Kragt and Jeffrey W Bennett. 2012. Attribute framing in choice experiments: how do attribute level descriptions affect value estimates? *Environmental and Resource Economics* 51, 1 (2012), 43–59.

[23] Jérôme Lang. 2007. Vote and Aggregation in Combinatorial Domains with Structured Preferences.. In *IJCAI*. 1366–1371.

[24] Jérôme Lang and Piotr Skowron. 2016. Multi-attribute Proportional Representation. In *AAAI*. 530–536.

[25] Jérôme Lang and Piotr Skowron. 2018. Multi-attribute proportional representation. *Artificial Intelligence* 263 (2018), 74–106.

[26] Jérôme Lang and Lirong Xia. 2009. Sequential composition of voting rules in multi-issue domains. *Mathematical social sciences* 57, 3 (2009), 304–324.

[27] Rob LeGrand, Evangelos Markakis, and Aranyak Mehta. 2007. Some results on approximating the minimax solution in approval voting.. In *AAMAS*, Vol. 7. 1185–1187.

[28] Minyi Li, Quoc Bao Vo, and Ryszard Kowalczyk. 2011. Majority-rule-based preference aggregation on multi-attribute domains with cp-nets. In *AAMAS-Volume 2*. 659–666.

[29] MPM Meuwissen, IA Van Der Lans, and RBM Huirne. 2007. Consumer preferences for pork supply chain attributes. *NJAS-Wageningen Journal of Life Sciences* 54, 3 (2007), 293–312.

[30] Toby Pereira. 2016. Proportional Approval Method using Squared loads, Approval removal and Coin-flip approval transformation (PAMSAC)-a new system of proportional representation using approval voting. *arXiv preprint arXiv:1602.05248* (2016).

[31] Luis Sánchez-Fernández, Edith Elkind, Martin Lackner, Norberto Fernández, Jesús A Fisteus, P Basanta Val, and Piotr Skowron. 2017. Proportional justified representation. In *AAAI*. 670–676.

[32] Ildikó Schlotter, Piotr Faliszewski, and Edith Elkind. 2011. Campaign Management under Approval-Driven Voting Rules.. In *AAAI*. 726–731.

[33] Piotr Skowron, Piotr Faliszewski, and Jerome Lang. 2016. Finding a collective set of items: From proportional multirepresentation to group recommendation. *Artificial Intelligence* 241 (2016), 191–216.

[34] Piotr Krzysztof Skowron and Piotr Faliszewski. 2015. Fully Proportional Representation with Approval Ballots: Approximating the MaxCover Problem with Bounded Frequencies in FPT Time.. In *AAAI*. 2124–2130.

[35] Srinivasaraghavan Sriram, Pradeep K Chintagunta, and Ramya Neelamegham. 2006. Effects of brand preference, product attributes, and marketing mix variables in technology product markets. *Marketing Science* 25, 5 (2006), 440–456.

[36] Alan D Taylor. 2002. The manipulability of voting systems. *The American Mathematical Monthly* 109, 4 (2002), 321–337.

[37] Lirong Xia, Vincent Conitzer, and Jérôme Lang. 2008. Voting on Multiattribute Domains with Cyclic Preferential Dependencies.. In *AAAI*. 202–207.

[38] Lirong Xia, Jérôme Lang, and Mingsheng Ying. 2007. Strongly decomposable voting rules on multiattribute domains. In *AAAI*. 776–781.