SHAPLEY VALUE APPROXIMATION FOR GAMES WITH DISTANT PLAYERS*

Ondřej Osička[†,1], Jaroslav Hrdina[1], Radovan Šomplák[2], Pavel Popela[1], and Martin Pavlas[2]

Brno University of Technology, Faculty of Mechanical Engineering
[1]Institute of Mathematics
[2]Institute of Process Engineering
Technická 2896/2, 616 69 Brno
Czech Republic
ondrej.osicka@vutbr.cz

Abstract: *The Shapley value provides a method for the division of total profit in cooperative games. Motivated by the waste management application, an algorithm to approximate the Shapley value is developed. The method is significantly faster than the classical approach to the Shapley value determination and can be used for any game where the distance of players can be measured and a critical value of distance, beyond which the cooperation between any two players is worthless, can be determined.*

Keywords: *cooperative game theory, Shapley value, approximation method, collaborative transportation, waste management*

1 Introduction

The Shapley value is a concept in cooperative game theory, introduced in [8], assigning to each player a portion of a profit obtained by a cooperation among all of them. According to [1], it is commonly used in collaborative transportation. In many applications, however, the Shapley value is computed only for games with few players. The computation time is growing significantly with more players cooperating. In real situations, such big coalitions might not be always easy, or even possible, to maintain, but the Shapley value can always serve as a benchmark for other solutions, showing the potential in cooperation.

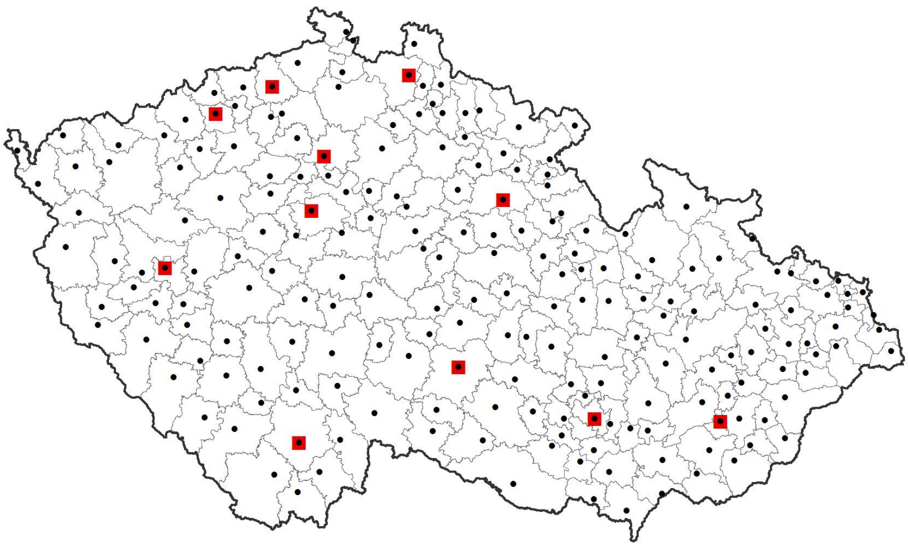

Fig. 1: Map of the administrative units (marked with a black dot) and waste incinerators (marked with a red square) for the waste management game in Czech Republic (source of spatial data: Arc ČR 500 v.3.2)

*This is a preprint of an article published in *Mendel*. Please cite the published version: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85014964155&partnerID=40&md5=42990e846db7ced34cc3d019182d5575

†Corresponding author. *E-mail address:* ondrej.osicka@vutbr.cz

In [5], a cooperative game of 206 players is modeled for a waste management problem. These players represent waste producers, administrative units of Czech Republic, which are choosing a way of waste disposal with costs as low as possible. A possible future situation defines their options as 11 waste incinerators in Czech Republic and 15 waste incinerators abroad. Their division within Czech Republic is illustrated in Fig. 1. A gate fee, the charge for waste processing, for each facility is known to all players. The case of a cooperation among all players was studied in [9] by the use of, there presented, NERUDA tool. Here, the cooperation is explored in more detail. The cooperation might become beneficial when the capacities of local waste incinerators are insufficient and the municipalities are forced to send their waste to distant ones. It is natural to assume that too distant players are unlikely to influence each other, hence any coalition of them seems worthless.

Here, an algorithm for a Shapley value approximation is developed. This algorithm can be used for any cooperative game where, for any two players, the efficiency of a cooperation between them can be predicted. This can be obviously said about games in which players are placed in a space and their cooperation is as effective as they are close to each other.

2 Basic Concepts

According to [6], an *n-player game in characteristic function form* is the pair (N, v) where $N = \{p_1, \dots, p_n\}$ is a set of n players and v is a real-valued function, defined on the subsets of N, satisfying conditions

$$v(\varnothing) = 0 \tag{1}$$

and

$$v(S \cup T) \geq v(S) + v(T) \qquad \text{if} \qquad S \cap T = \varnothing. \tag{2}$$

This function is denoted as the *characteristic function*. Any subset of N is called a *coalition*.

The *Shapley value* for an n-player game (N, v) is a vector $\varphi(N, v) = (\varphi_{p_1}(N, v), \dots, \varphi_{p_n}(N, v))$ defined by formula

$$\varphi_{p_i}(N, v) = \sum_{S \subseteq N \colon p_i \in S} \frac{(|S| - 1)!\,(|N| - |S|)!}{|N|!} \left(v(S) - v(S \backslash \{p_i\})\right). \tag{3}$$

Hence to determine the Shapley value, values of the characteristic function for all coalitions $S \subseteq N$ are needed. For an n-player game, it means the characteristic function values for 2^n coalitions. This is illustrated in Fig. 2.

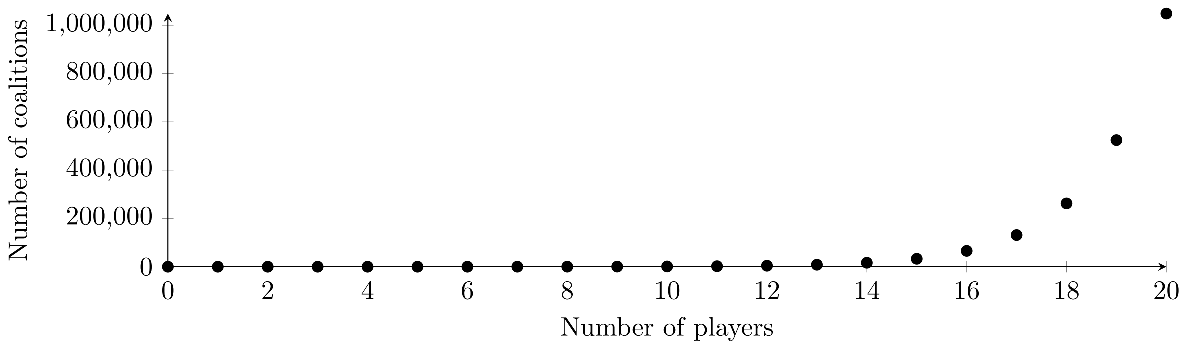

Fig. 2: Number of formable coalitions

For the waste management game of 206 players, the characteristic function values for approximately $1.03 \cdot 10^{62}$ coalitions are needed. Moreover, for this game, the characteristic function value is determined as a solution of a minimization problem. The combination of these makes the computation time very long. Omission of some coalitions would therefore be helpful.

3 Algorithm

This algorithm serves to compute a Shapley value approximation $\psi = (\psi_{p_1}, \dots, \psi_{p_n})$ for an n-player game (N, v). For this purpose, two inputs are needed. First of them is a critical distance, beyond which a cooperation between two players is expected to have no impact, and the other one is a maximum number of cooperating players, which is natural for the already mentioned reason that big coalitions are not always easy, or even possible, to maintain.

The characteristic function v must be in a form where $v(\{p_i\}) = 0$ for all $p_i \in N$. This prerequisite condition is not restrictive, because any characteristic function $\tilde{v}$ can be easily reformulated to this form by formula

$$v(S) = \tilde{v}(S) - \sum_{p_i \in S} \tilde{v}(\{p_i\}), \tag{4}$$

the computed approximation $\psi = (\psi_{p_1}, \ldots, \psi_{p_n})$ is then only modified by formula

$$\tilde{\psi}_{p_i} = \psi_{p_i} + \tilde{v}(\{p_i\}) \tag{5}$$

and $\tilde{\psi} = (\tilde{\psi}_{p_1}, \ldots, \tilde{\psi}_{p_n})$ is then the approximation for this game.

Step 1. Given a distance matrix $D = [d_{i,j}]$, where $d_{i,j}$ represents the distance between players p_i and p_j, and a critical distance d_{crit}, beyond which the cooperation is considered worthless, a matrix $A = [a_{i,j}]$ is created by formula

$$a_{i,j} = \begin{cases} 1 & \text{if} \quad d_{i,j} \leq d_{crit} \\ 0 & \text{otherwise} \end{cases} . \tag{6}$$

Step 2. Using the matrix A and given a maximum number of cooperating players c_{max}, a set of coalitions C is created. For this purpose, two approaches are used. A question rises, if a coalition $\{p_i, p_j, p_k\}$ should be included in this set when $a_{i,j} = 1$, $a_{j,k} = 1$, but $a_{i,k} = 0$. For a positive answer, Algorithm 2.1 is used. And for a negative one, Algorithm 2.2 is used.

Algorithm 2.1

if $c_{max} \geq 1$ **then**
  for $j = 1$ **to** c_{max} **do**
    set $C_j = \varnothing$
  end for
  for $i = 1$ **to** number of players **do**
    add $\{p_i\}$ to C_1
  end for
  for $j = 2$ **to** c_{max} **do**
    for all $S \in C_{j-1}$ **do**
      for $i = 1$ **to** number of players **do**
        if $p_i \notin S$ **and** $\sum_{k\colon p_k \in S} a_{i,k} \geq 1$ **then**
          add $S \cup \{p_i\}$ to C_j
        end if
      end for
    end for
  end for
  set $C = \bigcup_{j=1}^{c_{max}} C_j$
end if
add $\varnothing$ and N to C

Algorithm 2.2

if $c_{max} \geq 1$ **then**
  for $j = 1$ **to** c_{max} **do**
    set $C_j = \varnothing$
  end for
  for $i = 1$ **to** number of players **do**
    add $\{p_i\}$ to C_1
  end for
  for $j = 2$ **to** c_{max} **do**
    for all $S \in C_{j-1}$ **do**
      for $i = 1$ **to** number of players **do**
        if $p_i \notin S$ **and** $\prod_{k\colon p_k \in S} a_{i,k} = 1$ **then**
          add $S \cup \{p_i\}$ to C_j
        end if
      end for
    end for
  end for
  set $C = \bigcup_{j=1}^{c_{max}} C_j$
end if
add $\varnothing$ and N to C

Step 3. The value $\psi'(N, v, C) = (\psi'_{p_1}(N, v, C), \ldots, \psi'_{p_n}(N, v, C))$ is computed by formula

$$\psi'_{p_i}(N, v, C) = \sum_{S \in C\colon p_i \in S} \frac{(|S|-1)!\,(|N|-|S|)!}{|N|!} \left(v(S) - v'(S \backslash \{p_i\})\right), \tag{7}$$

where

$$v'(S \backslash \{p_i\}) = \begin{cases} v(S \backslash \{p_i\}) & \text{if} \quad S \backslash \{p_i\} \in C \\ v_{min}(S \backslash \{p_i\}) & \text{otherwise} \end{cases} , \tag{8}$$

where $v_{min}(S \backslash \{p_i\})$ is a solution of the following optimization problem. This approach is similar to the one used for a Shapley value refinement presented in [4]. Denoting $C = \{T_1, \ldots, T_{|C|}\}$ and $J = \{1, \ldots, |C|\}$, the integer programming problem looks as follows.

$$v_{min}(S\backslash\{p_i\}) = \min_{x_j:\, j\in J} \sum_{j\in J} v(T_j)x_j, \tag{9}$$

$$\text{s. t.} \quad \bigcup_{j\in J:\, x_j=1} T_j = S\backslash\{p_i\}, \tag{10}$$

$$\bigcap_{j\in J:\, x_j=1} T_j = \varnothing, \tag{11}$$

$$x_j \in \{0,1\} \qquad \forall j \in J. \tag{12}$$

In the case of a characteristic function not representing costs, but payoffs, the minimization should be replaced by a maximization.

Optimization problems of this type are commonly recognized as the assignment problems and occur also in many other applications like [3].

Step 4. The final step's only purpose is to preserve an assumption that the profit is completely divided among the players. Therefore, the final form of the Shapley value approximation here presented is a vector $\psi(N, v, C) = (\psi_{p_1}(N, v, C), \dots, \psi_{p_n}(N, v, C))$, where

$$\psi_{p_i}(N, v, C) = \psi'_{p_i}(N, v, C) + \frac{v(N) - \sum_{p_i\in N} \psi'_{p_i}(N, v, C)}{n}. \tag{13}$$

4 Computation Time

Computation tests were run for this algorithm. All computations were realized on the computer with Microsoft Windows 10 Home 64-bit, quad-core Intel Core i5-6300HQ at frequency 2.3 GHz and 8 GB of RAM. For complete input data, see [5].

Firstly, a classical approach for the Shapley value determination was implemented in MATLAB and run for the waste management game of less players. This game was determined by omission of other players. The computation time depends only negligibly on the choice of them, therefore, they are chosen randomly. Fig. 4 then illustrates the impact of the number of players on the time. According to this, the computation for a game of 206 players would take approximately $3.95 \cdot 10^{52}$ years.

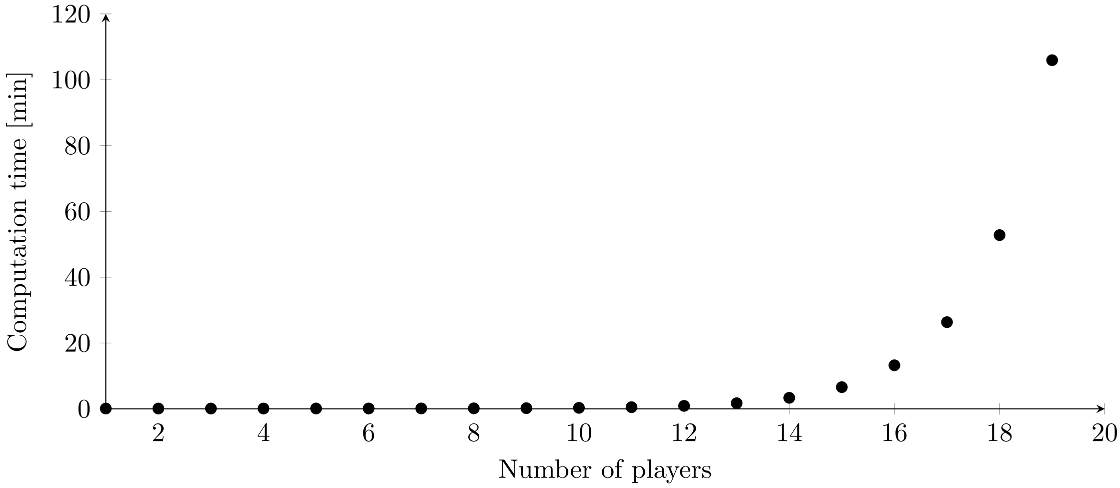

Fig. 4: Computation time of the classical Shapley value determination

Steps 1 and 2 of the presented algorithm were implemented in MS Excel and serve as input data for steps 3 and 4 implemented in MATLAB. The algorithm was run for the waste management game of 206 players. Computation times for multiple choices of d_{crit} and c_{max} are shown in Table 1 for the choice of algorithm 2.1 in step 2 and in Table 2 for the choice of algorithm 2.2. For the difference in results for some of the combinations, see Table 3.

For computation times of the above-mentioned NERUDA tool, see [7].

Table 1: Computation times of the algorithm with the choice of 2.1 (Combinations marked with '–' were unable to be computed due to insufficient memory of the MS Excel implementation.)

		c_{max} 5	6	7
d_{crit}	0	1 min 15 s	1 min 12 s	1 min 15 s
	10	1 min 17 s	1 min 18 s	1 min 18 s
	20	3 min 52 s	4 min 13 s	4 min 35 s
	30	38 min 3 s	1 h 30 min 39 s	4 h 23 min 9 s
	40	5 h 3 min 7 s	16 h 10 min 32 s	59 h 5 min 13 s
	50	24 h 4 min 47 s	–	–

Table 2: Computation times of the algorithm with the choice of 2.2

		c_{max} 5	10	15
d_{crit}	0	1 min 9 s	1 min 7 s	1 min 9 s
	10	1 min 11 s	1 min 11 s	1 min 11 s
	20	1 min 56 s	1 min 57 s	1 min 56 s
	30	4 min 23 s	4 min 22 s	4 min 23 s
	40	11 min 47 s	11 min 55 s	11 min 48 s
	50	38 min 17 s	45 min 47 s	45 min 11 s
	60	2 h 2 min 15 s	3 h 33 min 16 s	3 h 33 min 32 s
	70	5 h 54 min 33 s	19 h 20 min 55 s	19 h 42 min 2 s

5 Conclusion

This algorithm for the Shapley value approximation makes it possible to obtain a solution within a reasonable time. The accuracy of such solution depends mainly on the game itself. However, for games in which the threshold of beneficial coalitions cannot be determined, this approximation is useless.

Table 3 shows that, for the waste management game, it is not easy to choose the appropriate algorithm and set the exact threshold value, but the algorithm can be repeated until the result seems sufficient. It would be more complicated, if the model took into account the uncertainty on waste production. However, a prediction on the specific waste amount produced could be made from historical data or, for example, by pricing as studied in [2].

The algorithm could be yet extended by adding other conditions on reasonable coalitions. For example, assuming producers with large waste production being more likely worth cooperating with seems to be one of the possibilities for this extension.

Acknowledgement: The authors gratefully acknowledge financial support provided by the MEYS under the National Sustainability Programme I (Project LO1202) and financial support provided by Technology Agency of the Czech Republic within the research project no. TE02000236 "Waste-to-Energy (WtE) Competence Centre".

References

[1] Guajardo, M., Rönnqvist, M.: A review on cost allocation methods in collaborative transportation. In: International Transactions in Operational Research 23(3), 371–392 (2016). DOI 10.1111/itor.12205

[2] Hrabec, D., Popela, P., Roupec, J., Jindra, P., Novotný, J.: Hybrid algorithm for wait-and-see transportation network design problem with linear pricing. In: Proceedings of 21st International Conference on Soft Computing - MENDEL 2015, 183–188. Brno University of Technology, Brno (2015). ISSN 1803-3814

Table 3: A change in the value assigned to ten randomly chosen players (their total waste disposal costs) by using algorithms 2.1 and 2.2 and different values of c_{max} and d_{crit}

Algorithm	2.1	2.1	2.1	2.1	2.2	2.2	2.2
c_{max}	1	5	5	7	5	5	5
d_{crit}	0	20	40	20	20	40	70
Player 1	44,810,577	44,807,134	44,806,361	44,806,050	44,809,137	44,807,393	44,806,007
Player 2	283,698	283,929	284,065	284,036	283,805	283,971	283,933
Player 3	3,392,713	3,392,944	3,392,950	3,393,051	3,392,820	3,392,986	3,393,097
Player 4	772,546	772,776	772,900	772,884	772,652	772,818	772,859
Player 5	488,915	489,146	489,282	489,253	489,022	489,188	489,257
Player 6	1,056,360	1,056,591	1,056,704	1,056,698	1,056,467	1,056,632	1,056,641
Player 7	252,419	252,650	252,656	252,757	252,526	252,691	252,744
Player 8	309,399	309,630	309,731	309,737	309,506	309,671	309,754
Player 9	938,014	938,245	938,303	938,352	938,121	938,286	938,395
Player 10	602,500	602,731	602,819	602,838	602,607	602,773	602,883

[3] Kůdela, J., Popela, P.: Two-stage stochastic facility location problem: GA with Benders decomposition. In: Proceedings of 21st International Conference on Soft Computing – MENDEL 2015, 53–58. Brno University of Technology, Brno (2015). ISSN 1803-3814

[4] Myerson, R. B.: Graphs and Cooperation in Games. In: Mathematics of Operations Research 2(3), 225–229 (1977). DOI 10.1287/moor.2.3.225

[5] Osička, O.: Game Theory in Waste Management. Master's thesis, Brno University of Technology (2016)

[6] Owen, G.: Game Theory, fourth edn. Emerald, Bingley (2013). ISBN 978-1-78190-507-4

[7] Pavlas, M., Nevrlý, V., Popela, P., Šomplák, R.: Heuristic for generation of waste transportation test networks. In: Proceedings of 21st International Conference on Soft Computing – MENDEL 2015, 189–194. Brno University of Technology, Brno (2015). ISSN 1803-3814

[8] Shapley, L. S.: A Value for n-Person Games. In: Contributions to the Theory of Games, Volume II, 307–317. Princeton University Press, Princeton (1953). ISBN 978-0-691-07935-6

[9] Šomplák, R., Pavlas, M., Kropáč, J., Putna, O., Procházka, V.: Logistic model-based tool for policy-making towards sustainable waste management. In: Clean Technologies and Environmental Policy 16(7), 1275–1286 (2014). DOI 10.1007/s10098-014-0744-5