Scientific
Research

Interaction between Dutch Soccer Teams and Fans: A Mathematical Analysis through Cooperative Game Theory

Dongshuang Hou, Theo Driessen

Faculty of Electrical Engineering and Mathematics and Computer Science, University of Twente, Enschede, The Netherlands
Email: dshhou@126.com

ABSTRACT

Inspired by the first lustrum of the Club Positioning Matrix (CPM) for professional Dutch soccer teams, we model the interaction between soccer teams and their potential fans as a cooperative cost game based on the annual voluntary sponsorships of fans in order to validate their fan registration in a central database. We introduce a natural cost allocation to the soccer teams, based in a natural manner on the sponsorships of fans. The game theoretic approach is twofold. On the one hand, an appropriate cost game called "fan data cost game" is developed and on the other, it is shown that the former natural cost allocation agrees with the solution concept called "nucleolus" of the fan data cost game.

1. Club Positioning Matrix (CPM) of Professional Dutch Soccer

Five years CPM. The first lustrum of the (Dutch) "Eredivisie Effectenbeurs"is a fact. At the initiative of the professional teams in the Dutch soccer league called "Eredivisie", the first CPM research has been carried out October 2006 by one of the German leading research and consultancy companies in international sport business (*i.e.*, marketing and sponsoring) called "Sport + Markt" (www.sportundmarkt.com). October 2009 the fourth CPM research involved 4.500 participants randomly selected from the whole Dutch population (with the common feature to be a fan of soccer). Since these eighteen soccer teams were not satisfied at all by the point of time October, the fifth CPM research edition has been carried out among 4.500 participants in two stages, namely August 2010 at the beginning of the soccer season and January 2011 during the soccer winter break. The end of March 2011, the CPM 2011 scores have been sent to the professional Dutch soccer teams and published exclusively in the weekly Dutch soccer magazine "Voetbal International" (www.vi.nl) [1].

The CPM 2011 scores have direct consequences for the participating soccer teams since the allocation of media (television and broadcast) money among all the soccer teams is based equally on both the annual sport results and the average CPM scores over three years. The more CPM points, the more media money. During one half of a century, the annual sport results were dominated fully by the triple PSV Eindhoven (last Dutch championships in 2000, 2001, 2003, 2005, 2006, 2007, 2008), Ajax Amsterdam (2002, 2004, 2011), and Feyenoord Rotterdam (1974, 1984, 1993, 1999), with exceptions caused by DWS in 1964, AZ'67 Alkmaar in 1981 as well as 2009, and FC Twente Enschede in 2010. The top five of the last three annual sport results are listed in **Table 1**.

The annual CPM is a marketing instrument that measures the marketing value (through a professional jury of marketing specialists) as well as the imago of every professional soccer team (through the randomly selected soccer fans), which, in turn, is determined on the basis of six parts. Finally, the marketing value, the imago, and the annual sport result are put into some calculation model yielding the annual CPM scores.

The top five of the best marketing is as follows: 1) PSV; 2) SC Heerenveen; 3) FC Twente; 4) Ajax; 5) Feyenoord. Like the fourth edition, the team with the best imago is FC Twente due to its unique national championship, its successful participation in the international Champions League as well as the European League (till the quarter finales), and its new stadium called Grolsch Veste. FC Twente's imago is the best in the subfields attraction (charm), fascination, economical success, and the second best in the subfields emotional involvement and identification. The top six of the best imago is as follows: 1) FC Twente 699; 2) Ajax 641; 3) PSV 624; 4) SC Heerenveen 558; 5) FC Groningen 539; 6) Feyenoord 476.

In summary, the CPM score of FC Twente increased drastic, Feyenoord's score decreased drastic, so that the third ranking in the CPM 2011 scores is occupied by FC

R	2010-2011	Score	2009-2010	Score	2008-2009	Score
1	Ajax	73	FC Twente	86	AZ'67	80
2	FC Twente	71	Ajax	85	FC Twente	69
3	PSV	69	PSV	78	Ajax	68
4	AZ'67	59	Feyenoord	63	PSV	65
5	FC Groningen	57	AZ'67	62	SC Heerenveen	60

Twente. Ajax and PSV remain first and second due to the CPM results of the previous years. The top seven of the last three final CPM rankings are listed in **Table 2**.

Concerning the fanstatus, the CPM 2011 top five is as follows: 1) Ajax; 2) Feyenoord; 3) PSV; 4) FC Twente; 5) AZ'67. During the first CPM research October 2006, FC Twente started with a fanstatus of 250.000 fans, nowadays its fanstatus has been increased up to about 1.6 million, being the double of its previous edition.

2. The Fan Database Model

Given the current fan status as the model of the interaction between the professional Dutch soccer teams and their potential fans, our main goal is to apply the solution part of the mathematical field called "cooperative game theory". The so-called "players" are the soccer teams, each of which is endowed with a set of potential fans, each of which is supposed to validate its fan registration in a central database through an annual voluntary sponsorship to be cashed to the national soccer association. This annual sponsorship is said to be voluntary since it varies from fan to fan, each fan decides by him/herself about the contribution to be small or large. No registration if the potential fan is not willing to fulfill this sponsorship. In fact, any commitment to this sponsorship guarantees certain priorities to the fan, such as priority rights to purchase tickets for additional (inter)national soccer matches with or without discount, program booklets free of charge, and so on. Notice that any fan is allowed to be registered (in a central database run by the national soccer association) for a number of distinct soccer teams (not necessarily one team), while contributing the annual voluntary sponsorship once (at the beginning of the soccer season). **Table 3** surveys the essential notions about soccer teams and fans.

In summary, the fan database of professional Dutch soccer teams may be modeled as the triple $\left(N,\left(F_i\right)_{i\in N},\left(S_j\right)_{j\in F}\right)$ such that the "player set" N consists of the soccer teams, the set F_i consists of fans of soccer team i, and $S_j > 0$ represents the annual voluntary sponsorship of fan j. In fact, these sponsorships are combined to construct the following cost allocation

$y = \left(y_i\right)_{i\in N} \in R^n$. In the sequel let $|X|$ denote the cardinality of any finite set X. Consider the budget B as the sum of sponsorships of fans with a unique (unspecified) favourite soccer team (that is, $j \in F$ with $|N_j| = 1$). Firstly, factorize this budget in accordance with the appearance of the unique soccer team involved, that is $B = \sum_{i \in N} b_i$, with the understanding that $b_i = 0$ if there are no fans $j \in F$ with unique favourite soccer team i. Secondly, with reference to these factorizations, determine the deviations with respect to their average. In summary, charge to soccer team i the cost allocation amounting $y_i = \dfrac{B}{|N|} - b_i$ for all $i \in N$. In words, reward to soccer team i the negative amount $-b_i$, and charge the budget B equally among all the soccer teams. In particular, a soccer team i receives a reward (instead of a cost charge) if and only if the total sponsorship b_i exceeds the average $\dfrac{B}{|N|}$ of the budget. The larger b_i, the larger the reward to soccer team i. That is, soccer teams benefit from fans who are willing to contribute a large sponsorship. **Table 4** surveys the essential notions in the setting of cost allocations.

3. The Fan Data Cost Game

Our main goal is to support the cost allocation $\left(y_i\right)_{i\in N}$ from the viewpoint of cooperative game theory as the so-called nucleolus [2] of the suitably chosen fan data cost game $\langle N, c\rangle$. **Table 5** surveys the essential notions in the setting of fan data cost game.

Definition 3.1. The coalitional cost $c(S)$ of any non-empty coalition $S \subseteq N$ represents the loss (shortage) of coalitional sponsorship versus total sponsorship. That is, the fan data cost game $\langle N, c\rangle$ is given by

$$c(S) = \sum_{j \in F} s_j - \sum_{j \in F_S} s_j = \sum_{j \in F \setminus F_S} s_j \text{ for all } S \subseteq N, S \neq \phi \text{ (3.1)}$$

Note that the soccer teams are willing to cooperate to share the fan data information of the central database in order to solve the minimization problem of shortages of sponsorships in that $c(N) = 0$ reflecting the formation

Table 2. The top seven of the last three final CPM rankings is as follows (R = ranking).

R	CPM 2011	Score	CPM 2010	Score	CPM 2009	Score
1	Ajax	2.928	Ajax	2.888	Ajax	2.791
2	PSV	2.649	PSV	2.568	PSV	2.656
3	FC Twente	2.269	Feyenoord	2.237	Feyenoord	2.255
4	Feyenoord	2.199	AZ'67	2.106	AZ'67	2.165
5	AZ'67	2.086	FC Twente	2.059	SC Heerenveen	2.104
6	SC Heerenveen	2.071	SC Heerenveen	1.943	FC Groningen	1.804
7	FC Groningen	1.780	FC Groningen	1.552	FC Twente	1.661

Table 3. The essential notions about soccer teams and fans.

Symbol	Notation	Interpretation
N	$i \in N$	set of soccer teams; soccer team i called player
F_i	$j \in F_i$	set of fans of soccer team i; fan j of soccer team i
$F = \cup_{i \in N} F_i$	$j \in F_i$	set of all fans; fan j
$S_j > 0$		annual voluntary sponsorship of fan j in order to validate the fan registration
N_j	$\{i \in N \mid j \in F_i\}$	set of soccer teams of which j is a fan; equivalence: $i \in N_j \Leftrightarrow j \in F$

Table 4. The essential notions in the setting of cost allocations.

Formula	Interpretation
$B = \sum_{j \in F, |N_j| = 1} s_j$	sponsorship of fans with a unique unspecified favourite soccer team.
$b_i = \sum_{j \in F, N_j = \{i\}} s_j$	sponsorship of fans with the unique specified favourite soccer team $i, i \in N_i$
$y_i = \dfrac{B}{|N|} - b_i$	cost allocation charged to soccer team $i, i \in N$

Table 5. The essential notions in the setting of fan data cost game.

Formula	Interpretation
$i \in N$	Soccer team i called player.
$S \subseteq N$	Subset of soccer teams, called coalition.
$i \in S$	Soccer team i of coalition S.
$F_S = \cup_{i \in S} F_i$	Subset of fans of at least one soccer team of S.
$s_j > 0$	Annual voluntary sponsorship of fan j in order to validate the fan registration.
$\sum_{j \in F} s_j$	Total sponsorship of all fans.
$\sum_{j \in F_S} s_j$	Coalitional sponsorship of fans of at least one soccer team of S.
$\sum_{j \in F} s_j - \sum_{j \in F_S} s_j$	Coalitional loss (shortage) of sponsorship of soccer teams of coalition S
$c(S) = \sum_{j \in F} s_j - \sum_{j \in F_S} s_j$	Fan data cost of coalition S.
$c(N) = 0$	Motivation for cooperation to form the grand coalition N due to minimization of shortages of sponsorship.

of the grand coalition N. The next preliminary Lemma reports two essential properties of the fan data cost game $\langle N, c \rangle$.

Lemma 3.2. Let $\langle N, c \rangle$ be the fan data cost game of (3.1).

For all $S \subseteq N, S \neq N, S \neq \phi$,

$$c(S) \geq \sum_{i \in N \setminus S} b_i \quad \text{or equivalently, by (3.1)} \quad (3.2)$$

$$\sum_{j \in F \setminus F_S} s_j \geq \sum_{i \in N \setminus S} \sum_{j \in F, N_j = \{i\}} s_j \quad (3.3)$$

$$c(N \setminus \{i\}) = b_i \quad \text{for all} \quad i \in N. \quad (3.4)$$

Proof. (3.3) follows immediately from the inclusion $\{j \in F \mid N_j = \{i\}, i \in N \setminus S\} \subseteq F \setminus F_S$. In order to prove (3.4), fix player $i \in N$. By (3.1), we obtain the following chain of equations:

$$c(N \setminus \{i\}) = \sum_{j \in F \setminus F_{N \setminus \{i\}}} s_j = \sum_{j \in F, N_j = \{i\}} s_j = b_i.$$

Here the second equality is due to the following equivalences (given $i \in N$):

$$N_j = \{i\} \Leftrightarrow i \in N_j \quad \text{and} \quad k \notin N_j \quad \text{for all} \quad k \in N \setminus \{i\}$$

$$\Leftrightarrow j \in F_i \quad \text{and} \quad j \notin F_k \quad \text{for all} \quad k \in N \setminus \{i\}$$

$$\Leftrightarrow j \in F_i \quad \text{and} \quad j \notin F_{N \setminus \{i\}}$$

$$\Leftrightarrow j \in F \setminus F_{N \setminus \{i\}}.$$

4. The Nucleolus of the Fan Data Cost Game

Our study of the nucleolus involves the notion of excess $e(T, \mathbf{x}) = c(T) - \sum_{i \in T} x_i$, where $T \subseteq N, T \neq N, T \neq \phi$, and $\mathbf{x} = (x_i)_{i \in N} \in R^N$. It turns out that the level of the smallest excesses with respect to our cost allocation $\mathbf{y} = (y_i)_{i \in N}$ is composed of all the $(|N| - 1)$-person coalitions $N \setminus \{i\}, i \in N$. Indeed, on the one hand, by (3.4), and on the other, by (3.2), it holds for all $i \in N$ and all $T \subseteq N, T \neq N, T \neq \phi$.

$$e(N \setminus \{i\}, \mathbf{y}) = c(N \setminus \{i\}) - \sum_{k \in N} y_k + y_i = b_i + y_i = \frac{B}{|N|}$$

and further

$$e(T, \mathbf{y}) = c(T) - \sum_{i \in T} y_i = c(T) + \sum_{i \in T} b_i - \frac{B}{|N|} \cdot |T|$$

$$\geq \sum_{i \in N} b_i - \frac{B}{|N|} \cdot |T| = B - \frac{B}{|N|} \cdot |T| \geq \frac{B}{|N|}$$

Hence, all $(|N| - 1)$-person coalitions have the smallest excess among non-trivial coalitions with respect to our cost allocation $\mathbf{y}$ and according to Kohlberg's criterion [3], this suffices to conclude that our cost allocation $\mathbf{y}$ agrees with the nucleolus of the fan data cost game $\langle N, c \rangle$ of the form (3.1). Thus, in the setting of fan data

cost games, its nucleolus is fully solved by the explicit form of our cost allocation $y_i = \frac{B}{|N|} - b_i, i \in N$, and so, numerical methods and computational complexity are replaced by theoretical results due to the significant property (3.2) of the fan data cost game. Clearly, the number of constraints in (3.2) is the same as the exponential number of non-trivial coalitions. According to the equivalent property (3.3), slight changes in the sponsorships do not affect the strict inequalities and so, stability applies to some extent.

5. A Four-Person Example of a Fan Data Cost Game

Example 5.1. The four soccer teams of Ajax Amsterdam (A), Olympique Lyon (L), Real Madrid (M), and Dynamo Zagreb (Z) compete against each other within one group during the first round of the Champions League 2011-2012. Suppose that six television stations are appointed to broadcast the mutual matches such that the public Spanish TV station 3 is interested in all the four soccer teams, the public French TV station 2 is interested in every team except Zagreb, the national Dutch and Croatian TV stations 1 and 6 respectively only in their home team, similar to the local French and Spanish TV stations 4 and 5 respectively. That is, the data sets F_A, F_L, F_M, F_Z of fans of these four soccer teams are given by $F_A = \{1, 2, 3\}$, $F_L = \{2, 3, 4\}$, $F_M = \{2, 3, 5\}$, $F_Z = \{3, 6\}$. Based on the existence of the TV stations 1, 4, 5, 6 with a unique specified favourite soccer team A, L, M, Z, the sponsorships s_1, s_4, s_5, s_6 by these four TV stations are allocated to the corresponding soccer teams as separable benefits b_A, b_L, b_M, b_Z. In the second stage the total budget $B = b_A + b_L + b_M + b_Z = s_1 + s_4 + s_5 + s_6$ is divided equally among all the four soccer teams resulting in the allocations amounting

$$y_A = \frac{B}{4} - s_1, \quad y_L = \frac{B}{4} - s_4, \quad y_M = \frac{B}{4} - s_5 \quad \text{and}$$

$$y_Z = \frac{B}{4} - s_6.$$

This final allocation is supported by the game theoretic approach as the solution concept called nucleolus of the fan data cost game $\langle N, c \rangle$ listed in **Table 6**. Notice that the fourth column concerning the cost of any coalition is equal to the fifth column concerning the sum of separable benefits of the soccer teams in the complementary coalition, except for coalition $\{Z\}$. The dominance of the fourth column to the fifth column is the most significant property (3.2) of the fan data cost game.

The core of the fan data cost game is a quadrilateral with extreme points $(B - s_1, -s_4, -s_5, -s_6)$,

AM

Table 6. The essential notions in the example.

$S \subseteq N$	F_S	$F \setminus F_S$	$c(S) = \sum_{j \in F \setminus F_S} s_j$	$\sum_{i \in N \setminus S} b_i$
{A}	{1,2,3}	{4,5,6}	$s_4 + s_5 + s_6$	$b_L + b_M + b_Z$
{L}	{2,3,4}	{1,5,6}	$s_1 + s_5 + s_6$	$b_A + b_M + b_Z$
{M}	{2,3,5}	{1,4,6}	$s_1 + s_4 + s_6$	$b_A + b_L + b_Z$
{Z}	{3,6}	{1,2,4,5}	$s_1 + s_2 + s_4 + s_5$	$b_A + b_L + b_M$
{A,L}	{1,2,3,4}	{5,6}	$s_5 + s_6$	$b_M + b_Z$
{A,M}	{1,2,3,5}	{4,6}	$s_4 + s_6$	$b_L + b_Z$
{A,Z}	{1,2,3,6}	{4,5}	$s_4 + s_5$	$b_L + b_M$
{L,M}	{2,3,4,5}	{1,6}	$s_1 + s_6$	$b_A + b_Z$
{L,Z}	{2,3,4,6}	{1,5}	$s_1 + s_5$	$b_A + b_M$
{M,Z}	{2,3,5,6}	{1,4}	$s_1 + s_4$	$b_A + b_L$
{A,L,M}	{1,2,3,4,5}	{6}	s_6	b_Z
{A,L,Z}	{1,2,3,4,6}	{5}	s_5	b_M
{A,M,Z}	{1,2,3,5,6}	{4}	s_4	b_L
{L,M,Z}	{2,3,4,5,6}	{1}	s_1	b_A
{A,L,M,Z}	{1,2,3,4,5,6}	ϕ	0	0

$\left(-s_1, B - s_4, -s_5, -s_6\right)$, $\left(-s_1, -s_4, B - s_5, -s_6\right)$ and $\left(-s_1, -s_4, -s_5, B - s_6\right)$, where $B = s_1 + s_4 + s_5 + s_6$. Without going into details, we report that the nucleolus of any fan data cost game coincides with the center of gravity of the core. In case a seventh TV station joins with only interest in the Dutch team Ajax, then the separable benefit of Ajax increases with the sponsorship s_7 of the new TV station up to $s_1 + s_7$ whereas the final cost charge to Ajax lowers by $3 \cdot s_7 / 4$, that is the cost charge to any of the other three soccer teams increases with $s_7 / 4$.

6. A Remark about the Core of the Sponsorship Game

In the setting of the division problem of the total budget B among the soccer teams, it is natural to study the sponsorship game $\langle N, v \rangle$ defined by

$$v(S) = \sum_{j \in F_S} s_j \quad \text{for all} \quad S \subseteq N, S \neq \phi \qquad (6.1)$$

Clearly, straight from its definition, $v(N) \geq B$. Unfortunately, this sponsorship game has the drawback that its so-called "core" is empty. For that purpose, consider the set of "reasonable" payoff vectors of the game consisting of efficient payoff vectors $x = \left(x_i\right)_{i \in N} \in R^N$ satisfying $\sum_{i \in N} x_i = v(N)$ as well as lower and upper bounds for the individual payoffs such that $v(\{i\}) \leq x_i \leq v(N) - v(N \setminus \{i\})$ for all $i \in N$. In the context of the sponsorship game $\langle N, v \rangle$ of the form

(6.1), it holds for all $i \in N$

$$v(N) - v(N \setminus \{i\}) = \sum_{j \in F} s_j - \sum_{j \in F_{N \setminus \{i\}}} s_j = \sum_{j \in F \setminus F_{N \setminus \{i\}}} s_j$$
$$= \sum_{j \in F, N_j = \{i\}} s_j = b_i$$

So any reasonable payoff vector x satisfies $x_i \leq b_i$ for all $i \in N$ and consequently, by summing up, $v(N) \leq B$. This contradicts the earlier observation $v(N) \geq B$. So, reasonable payoff vectors do not exist, and hence, the core of the sponsorship game is empty too (as a subset).

For an introductory book on cooperative game theory, we refer to [4]. For a paper with similar contents, we refer to [5].

7. Acknowledgements

REFERENCES

[1]   Van Duren and T. Knipping, "Het Ijzeren Gordijn Is Gesloopt," *Voetbal International* (*Weekly Soccer Magazine, in Dutch*), Vol. 46, No. 13, 30 March 2011, pp. 116-121.

[2]   D. Schmeidler, "The Nucleolus of a Characteristic Function Game," *SIAM Journal of Applied Mathematics*, Vol.

17, No. 6, 1969, pp. 1163-1170. doi:10.1137/0117107

[3]   E. Kohlberg, "On the Nucleolus of a Characteristic Func-tion Game," *SIAM Journal of Applied Mathematics*, Vol. 20, No. 1, 1971, pp. 62-66.

[4]   T. S. H. Driessen, "Cooperative Games, Solutions, and Applications," Kluwer Academic Publishers, Dordrecht, 1988.

[5]   P. Dehez and D. Tellone, "Data Games: Sharing Public Goods with Exclusion," *Journal of Public Economic The-ory*, 2012.

AM