International Journal of Information Technology & Decision Making
Vol. 15, No. 4 (2016) 763–789

DOI: 10.1142/S0219622016400022

Expert System for Ice Hockey Game Prediction: Data Mining with Human Judgment

Wei Gu

Donlinks School of Economics and Management
University of Science and Technology Beijing
Beijing 100083, China
guwei@ustb.edu.cn

Thomas L. Saaty

Joseph M. Katz Graduate School of Business
University of Pittsburgh
Pittsburgh 15206, PA, USA
saaty@katz.pitt.edu

Rozann Whitaker

Creative Decisions Foundation
Pittsburgh 15213, PA, USA
rozann@creativedecisions.net

Published 25 July 2016

This paper describes an expert system to predict National Hockey League (NHL) game outcome. A new method based on both data and judgments is used to estimate the hockey game performance. There are many facts and judgments that could influence an outcome. We employed the support vector machine to determine the importance of these factors before we incorporate them into the prediction system. Our system combines data and judgments and used them to predict the win–lose outcome of all the 89 post-season games before they took place. The accuracy of our prediction with the combined factors was 77.5%. This is to date the best accuracy reported of hockey games prediction.

Keywords: Prediction; expert system; hockey game; data analysis; judgment.

1. Introduction

Does investigating the past aid foretelling the future? Bohr[1] has fittingly put it "Prediction is very difficult especially with regard to the future". Indeed, predicting the outcome of sports competition is very difficult, but it has been given close and extensive attention for a long time. Research on sports competition has evolved into a serious field of investigation whose major goal is to achieve high accuracy. The outcome of a sports game is unknown until the last "second" of play. There are many primary factors that influence the outcome: luck, psychological factors, refereeing, the strength of a team, etc. Most factors are uncertain as to how much they influence the outcome of the game, which makes the outcome harder to predict. It is commonly believed that group sports such as basketball, football, and hockey are much harder to predict than individual sports (e.g., badminton, ping pong and tennis). The uncertainty of the outcome of competitive games increases the attractiveness and excitement of sports competition.

Different from basketball (NBA) and football (NFL), on which considerable research has been done over the years, research on hockey has been sparse.[2] Weissbock *et al.* found no previous work in the machine learning (ML) community has predicted the winner in a hockey game. There are more uncertain factors in hockey than in other games, which make ice hockey harder to analyze and to forecast. This is because (i) more players play for a short while and rest for a short while, (ii) higher speed, (iii) collisions and combat being allowed and (iv) the desultory nature of play. Except for the work by Saaty and Zhang[3] on basketball, the literature on sports has relied solely on past data to predict the outcome of future games. No judgments about intangible factors were quantified in a meaningful way for use in prediction.

In this research, we developed an expert system to predict the outcome of the National Hockey League (NHL) game. The system is developed based on 1,230 games (2,460 records) results, standings, statistics and match details, with equal emphasis on judgments and experiences of the experts. We aim to create a comprehensive and highly accurate system to forecast hockey game performance.

The uniqueness of this paper is fourfold:

(a) We are the first to predict hockey game outcome using a method integrating both the data and judgments.
(b) The factors adopted for prediction are validated for their usefulness in producing an accurate outcome based on historical data analysis.
(c) The data from the data analysis are used to form expert judgments alongside with intangible, knowledge and experience.
(d) The accuracy of our prediction with the combined factors was 77.5% which to the best of our knowledge is the best accuracy of hockey game predictions.

The rest of the paper is organized as follows: Sec. 2 reviews related research, including sports prediction accuracy. In Sec. 3, we identify the key factors influencing the outcome of ice hockey games. Section 4 details our expert system of predictions, which incorporates both data and judgments. We describe in Sec. 5 our predictive results, and compare them with other methods commonly used in other sports (e.g., basketball and soccer). Finally, in Sec. 6, we discuss the effectiveness of this approach and conclude this research.

2. Literature Review

2.1. *Methods for sports game prediction*

Only a handful of researchers have endeavored to predict hockey games. Voyer & Wright,[4] used regression analysis to estimate the relationships among the variables. Regression analysis helps one understand how the expected value of the dependent variable changes when the independent variable varies while holding other independent variables constant. They considered players' performance in shooting, scoring goals, getting the puck, etc. However, they did not consider the interactions and feedback of factors.

Weissbock *et al.*[2,5] and Weissbock and Inkpen[6] from University of Ottawa have studied the NHL teams. They applied ML to predict hockey game performance by jointly considering previous games, pre-game textual reports and opinions of pre-game reports. They contribute to the application of ML to game prediction. However, their best accuracy is only at 62%, an upper bound limited by the nature of ML in this application.

Pischedda (2014) extended Ottawa's data to continuous and categorical classes, and use his model for real life betting. The model heavily depends on commentators' opinions. Alternatively, Depken *et al.*[7] and Kolev *et al.*[8] predicted the NHL regular season game using shootouts; while Morgan *et al.*[9] used the decision tree to predict with 64.3% accuracy.

Neural networks is another technique widely used to predict the outcome of sports games,[10] such as soccer,[11] NFL,[12,25] and basketball (NBA).[13,14] The reported predictive accuracy of the neural network approach is in the range of 74–78%. Finally, Zimmermann *et al.*[15] compared various ML techniques, and concluded Bayesian theory outperforms other procedures (e.g., decision tree, neural networks).

In summary, the conventional regression model is ill-suited for subjective judgments. On the other hand, ML is the most popular method. Yet, it relies heavily on subjective opinions and textual contents, which often results in low prediction accuracy. In this research, we integrate expert judgment and information derived from historical data to construct the expert system for ice hockey game predictions. We comprehensive consider both the tangible and intangible factors affecting the game outcomes and take experts knowledge and experience into account. The prediction accuracy of the proposed expert system is not limited by the upper bound inherent in the ML method.

2.2. *Factors influencing the outcome of sports game*

Numerous factors could influence sports game's outcome. In identifying factors, Feltz and Lirgg,[16] focused on team-based model. They used efficacy, belief and judgment to predict hockey players' performance. They find that if a team wins, its efficiency improves; otherwise its efficiency drops. Players' effectiveness, however, is not affected by win or loss.

To help with hiring decision, Perlini and Halverson,[17] examined criteria such as size/strength, skating/speed/power of stride, shot/scoring, etc. They found emotional intelligence, intrapersonal competency, and general mood have an equally significant influence on the outcome of a game. On the other hand, Buttrey *et al.* (2011) found that goals in NHL teams depend on the strength of opponents, the home-ice advantage, and players' skills and dexterity. Macdonald,[19] proposed new statistics such as face-offs and hits as predictors of goal scoring; they also found that many predictor variables are highly correlated.

Then again, Saaty and Zhang[3] propose an integrated method using both probabilities (Bayes theory) and judgments, to deal with prediction; and apply it to NBA basketball games prediction. Unlike basketball and football, where players are positioned in pair, in hockey each player in the team offends and defends simultaneously while collaborating with each other. Also, in the post season of the hockey game, the game performances are sequentially dependent. Therefore, it is important to identify an effective way to address interdependence when predicting collaborative sports games.

To address the need of hockey game prediction, we developed an expert system based on both the historical team performances and expert judgments. Through the inclusions of the subjective and objective information in the analytic network process (ANP) framework, we allow for the inclusion of interdependence and feedback in the model to tackle the intricacy of competitive NHL hockey games.

3. Preparation of Diverse Information for the Expert System

There are many available data that can be collected for hockey games. A performance metric is a measure that shows a team's performance and accomplishment. These numbers come in many forms: percentages, goals scored, shots on goal, etc. These metrics is complicated as sometimes larger numbers are better (Goals scored), while other times smaller numbers are better (Goals against). The performance metrics in this research is a team metrics, that in some cases have to be aggregated from individual player's metrics (e.g., shots on goal, faceoff wins, or blocked shots). These are numbers obtained by counting or combining counts or statistics. Some metrics are not performance related, they are simply categorical data about the game being played (e.g., home versus away game; regular versus post season; referees' names; coaches' names; the number of attendees; whether teams change coaches in the course of the season). Others are compound metrics obtained by combining simple metrics. For example, Corsi counts different kinds of shot attempts and is the sum of on-goal shots, missed shots and blocked shots.

Figure 1, presents the framework of this research and the schema of the expert system, we developed for hockey game prediction. In the following, we detail each step of the expert system, which corresponds to the section numbers below.

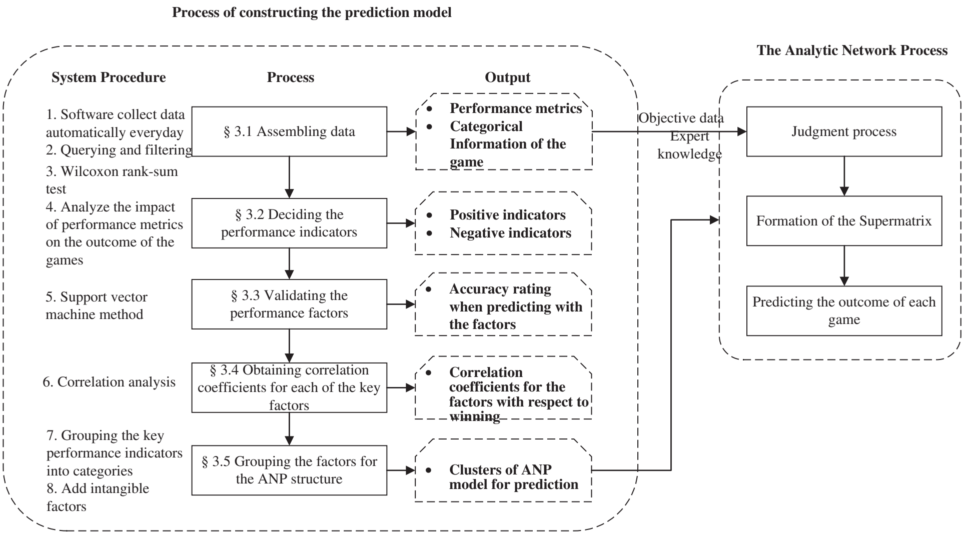

Fig. 1. The framework of expert system of the ice hockey game prediction.

3.1. *Assembling the data*

There were 1,230 regular season games in the 2014–2015 season, equivalent to 2460 sets of metrics, as it takes two teams to play in each game. The metrics for each team in a game constitute a record. Many data overlapped in different websites, and data often are presented in different formats. To filter/manage data so as to ensure compatibility, we developed a software program to automate the data search and collection process. When overlapping occurs (e.g., Corsi metric containing several overlapped info), we have to eliminate the overlapped metrics through additional procedure. We also eliminated shootout metrics because they result from very different game-time situations that occur erratically. The metrics for the penalty times was included, but not given special treatment. We ended up with 30 metrics (factors) shown in Table 1, of which the first four are observational factors and the remaining 24 are performance metrics. A sample of raw material teams records is shown in Table 2.

3.2. *Reducing the 24 performance metrics to 17 key performance metrics*

We used the Wilcoxon rank-sum test to check the significance of each individual factor. For each performance metric, we divide the records into two groups: win and loss, then we use the rank-sum test to compare each performance metrics from all games won to those lost. In the rank-sum test we used 0 and 1 to indicate win and loss respectively for each of the two teams involved in a game. We found that seven of the performance metrics are not significant in predicting the game performance. The 17

Table 1. Thirty situational and performance data factors.

Item	Explanation	Item	Explanation
Team	Team name	BSF	Blocked shots for
Season	Season type	BSA	Blocked shots against
R/H	Road game or home game	SF%	Shots on goal for percentage
Date	Game date	SF	Shots on goal for total
Dec	Decision of the game	SA	Shoot on goal against total
SV%	Save percentage	Shoot%	Shooting percentage
CF%	Corsi%: The percentage of on-ice shot attempts (on goal, missed, or blocked) versus shot successes (goals)	FO%	Faceoff winning percentage
CF	Corsi for total	FO_W	Faceoff won
CA	Corsi against total	FO_L	Faceoff lost
FF%	Fenwick for percentage of total	HIT	Hits
FF	Fenwick for total	HIT-	Hits taken
FA	Fenwick against total	PN	Penalties
MSF	Missed shots for	PN-	Penalties drawn
MSA	Missed shots against	Pen-D	Penalties differential

factors that were found to be significant in predicting the game results are shown in Table 3 as TRUE. The "+" sign means the indicator is the more the better (the save percentage), The "−"sign means the indicator is the less the better (the shots against).

Table 2. Sample of raw data team records for the 24 performance metrics.

Team	VAN	MTL	S.J	PHI	BOS	L.A	TOR	CGY
Road/ Home	R	R	R	R	H	H	H	H
Date	2014/10/8	2014/10/8	2014/10/8	2014/10/8	2014/10/8	2014/10/8	2014/10/8	2014/10/8
Dec	W	W	W	L	W	L	L	L
Sv%	92	88.9	100	93.9	95	86.7	87.5	87.9
CF%	52.6	50.4	46	43.1	56.9	54	49.6	47.4
MSF	6	15	15	9	10	19	10	7
MSA	7	10	19	10	9	15	15	6
FF%	54.9	56	45.9	40.3	59.7	54.1	44	2
FF	39	47	45	29	43	53	37	4
FA	32	37	53	43	29	45	47	L
SF%	56.9	54.2	46.9	37.7	62.3	53.1	45.8	43.1
SF	33	32	30	20	33	34	27	25
SA	25	27	34	33	20	30	32	33
FO%	38.6	58.7	59.2	41	59	40.8	41.3	61.4
FO_W	27	37	42	25	36	29	26	43
FO_L	43	26	29	36	25	42	37	27
BSF	11	11	13	15	15	15	20	13
BSA	13	20	15	15	15	13	11	11
Sht%	12.1	12.5	13.3	5	6.1	0	11.1	8
HIT	16	21	22	26	23	30	35	32
HIT-	32	35	30	23	26	22	21	16
PN	6	2	6	4	6	5	2	5
PN-	5	2	5	6	4	6	2	6
PenD	−1	0	−1	2	−2	1	0	1

Table 3. Results from the Wilcoxon rank-sum test.

Factors	Direction	Whether affect	Factors	Direction	Whether affect
G+/−	+	TRUE	SF	+	TRUE
Save%	+	TRUE	SA	−	TRUE
CF%	−	TRUE	Shoot%	+	TRUE
CF	−	TRUE	FO%	+	TRUE
CA	+	TRUE	FO_W	+	FALSE
FF%	+	FALSE	FO_L	−	FALSE
FF	+	FALSE	HIT	−	TRUE
FA	−	FALSE	HIT-	+	TRUE
MSF	−	TRUE	PN	−	FALSE
MSA	+	TRUE	PN-	+	FALSE
BSF	−	TRUE	Pen-D	+	TRUE
SF%	+	TRUE	BSA	+	TRUE

3.3. *Validating the 17 key performance metrics*

We now employed a ML technique, support vector machine (SVM) to validate the usefulness of the 17 variables (performance metrics) selected above. SVM is a discriminative classifier using a supervised ML model for pattern recognition, classification, and prediction. We use the 17 "important" game performance metrics derived in §3.2 as inputs to the SVM (see Table 3).

We built an SVM classifier from the historical game data set $S = \{(x^i, y^i), i = 1, 2, \ldots, m\}$, $X^i = \left(x_1^i, x_2^i, \ldots, x_n^i\right)$ where m represents the number of instances ($m = 1230$), and n is the number of input performance metrics ($n = 17$). z is a linear combination of the attributes (x_i) multiplied by corresponding weights (ω_i), plus a noise term (b).

$$Z = \omega_1 x_1 + \omega_2 x_2 + \cdots + \omega_n x_n + b, \tag{1}$$

where X_i denotes the inputs of the 17 performance metrics for game i. $y = -1$ indicates a loss, while $y = 1$ indicates a win. Eq. (1) can be determined by the sequential minimal optimization (SMO) algorithm (Platt, 1998), the fastest approach to solve SVM. If $Z > 0$, the analyzed team is expected to win the game; otherwise, a loss is expected for that team.

The training set is comprised of 1,230 randomly selected game records, and the remaining 1,230 records constitute the test set. We found the resulting trained SVM classifier can correctly classify 88.3% of the test data set. Thus, we conclude that the performance metrics (factors) we selected are valuable in predicting NHL games.

3.4. *Obtaining correlation coefficients for each of the key factors*

We then identify the correlations between the win/loss outcomes and each of the 17 metrics for the 2,460 regular season game records (Table 4).

When several factors are found to be significantly correlated with the response variable, it is necessary to use correlation analysis to understand their linear associations.

Table 4. Correlation coefficients for the factors with respect to winning.

Factors	Correlation	Factors	Correlation
G+/−	0.803	SF%	0.103
Save%	0.566	SF	−0.103
CF%	0.027	SA	−0.163
CF	−0.076	Shoot%	0.229
CA	−0.112	FO%	0.048
MSF	−0.022	HIT	−0.090
MSA	−0.041	HIT-	0.000
BSF	−0.136	Pen D	0.077
BSA	−0.005		

The correlation coefficients is shown in Table 4. We used these coefficients from the regular season as references for experts when employing the proposed expert system.

3.5. *Grouping the factors for an ANP structure*

We structure an ANP model by first grouping the key performance indicators into categories. Next, we add intangible factors that are not from data collected earlier, as they are important judgments from experts.

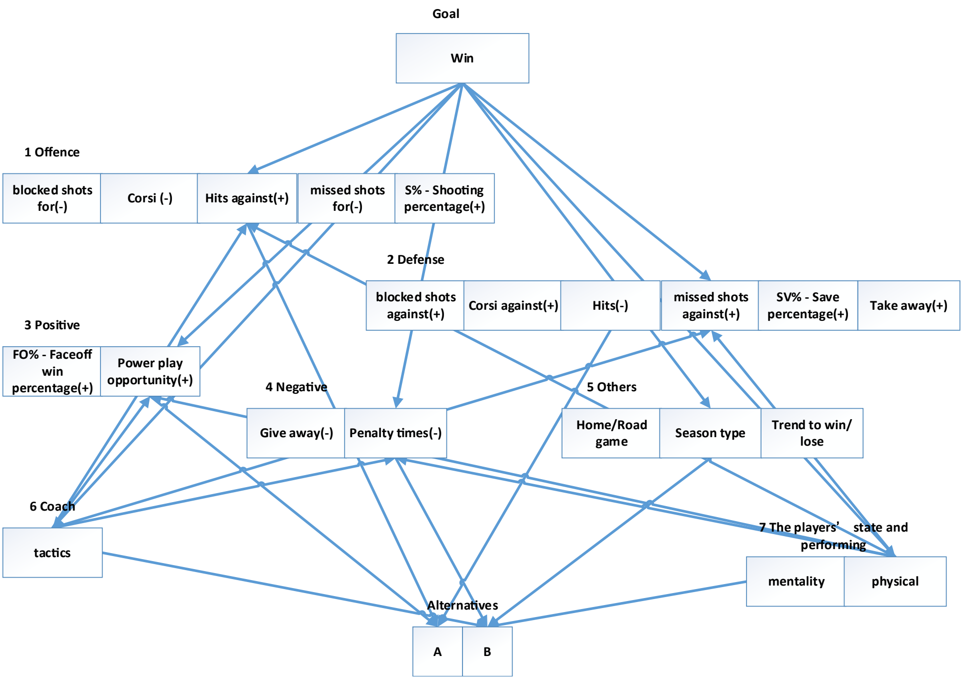

Fig. 2. The ANP network for predicting NHL playoff matches.

Table 5. Clusters for prediction.

Clusters	Factors	Type
Offense	Corsi* (−)	Data
	Missed shots for(−)	Data
	Blocked shots for(−)	Data
	Hits against(+)	Data
	S% - Shooting percentage(+)	Data
Defense	Corsi against(+)	Data
	Missed shots against(+)	Data
	Blocked shots against(+)	Data
	Hits(−)	Data
	Take away(+)	Data
	SV% - Save percentage(+)	Data
Positive	Power play opportunity(+)	Data
	FO% - Faceoff win percentage(+)	Data
Negative	Penalty times(−)	Data
	Give away(−)	Data
Others	Season type	Fact
	Home/Road game	Fact
	Trend is winning or losing	Judgment
Coach(tactics)	Tactics	Judgment
The player's state and performance	Includes the state of physical and mental readiness	Judgment

Notes: Corsi* means on-ice shot attempts (on goal, missed, or blocked).

The factors of Table 3 are grouped into four clusters: (1) Offense, (2) Defense, (3) Positive and (4) Negative (see Fig. 2). In addition, (5) situational indicators (Home/Road game, season type and trend to win/lose) form the fifth cluster named *Others* in Fig. 2. Finally, (6) Coach as well as (7) Players' mental/physical states, which are subjectively judged, constitute the last two clusters of Fig. 2.

The seven clusters in all contain 20 indicators (see Table 5). Offense comprises factors that are indicative of a team's strength in attack. Defense includes factors that signify a team's ability in guarding against the opponent. Table 5 shows that the first four clusters employ the information from data, the *Others* cluster contains Fact, while the rest clusters use judgments.

4. Model of the System-the Analytic Network Process used for Prediction

Each regulation game of the NHL is played between two teams for 60 min consisting of three 20 min periods. The international ice hockey federation (IIHF) also has postgame rules for how to handle ties. At the end of the 60 min of regulation time, the team with the most goals wins the game. If a game is tied after regulation time is over, overtime rules ensue. During the regular season, there is overtime play that lasts for 5 min with four-players (plus the goalie) on each side and the game terminates when one team scores (known as sudden death). If there is no score, the game enters a shootout in which a player from one team gets one shot at the goal which is defended by the goalie of the other team. Then a player from the other team gets a shot at the goal defended by the goalie of the first team. This continues for three rounds until one team scores more goals than the other in a shootout round and wins the game. If tied, they continue until one of them scores a goal and the other does not. The team with the most goals during a three-round shootout wins the game. If the game is still tied after the three shootout rounds, the shootout continues but becomes sudden-death with the first team to score thereby winning the game. There are no shootouts during the Playoffs. Instead, multiple sudden-death, 20 min five-on-five periods are played until one team scores.

The NHL is a professional ice hockey league composed of 30 member clubs: 23 in the United States and 7 in Canada. Teams play a total of 82 games each regular season from October to April for a total of 1,230 games. There are 2 conferences with three divisions with five teams belonging to each division. Within each division, the teams confront each other six times a year with a total of 24 games for each team within the division; within the conference each team confronts 10 teams outside its division but in its conference only four times. In the remaining 18 games, it plays at least once each of the 15 teams in the other conference and three of them twice bringing the total to 82. At the end of a regular season, the top eight teams from each conference qualify for the next round competition. The eventual winner wins four best-of-seven series in an elimination tournament and becomes the Stanley Cup Champion.

The Analytic Hierarchy Process (AHP)[20] developed by Saaty[20] is a new scientific decision method based on hierarchical structures and making judgments. Saaty[21] elaborated it into the ANP,[21,22] based on network structures with dependence and feedback. In the ANP, networks of clusters of elements are used instead of the hierarchic levels of elements of the AHP. An ANP model can offer a solution to complex multi-criteria problems that have little objective supporting data.[23] Thus, it is thus appropriate to analyze a complex problem such as predicting team sports game outcomes.

To predict the winner of a match before the game, an expert (or experts) answers questions about the relative importance of the indicators and about the relative performance of the teams based on historical data. Using judgments, we can calculate the influence priority of the elements in an alternative cluster on the elements in a criterion cluster with respect to each control criterion and predict who will win in each match. Thus, the ANP can be used as a prediction model in the expert system for ice hockey games.

We combine judgments and data in an ANP model to predict the outcome of each of the 89 playoff games in the NHL 2014–2015 season. We use the same ANP structure for every game but customize it with data for the two teams playing. The process is explained in the following section.

The fundamental scale used for judgments in the ANP is shown in Table 6. A number from the fundamental scale is chosen to represent the intensity of the

Table 6. Fundamental scale Saaty, 1982.

1	Equal importance
3	Moderate importance of one over another
5	Strong or essential importance
7	Very strong or demonstrated importance
9	Extreme importance
2,4,6,8	Intermediate values
	Use reciprocals for inverse comparisons

relationship between the factors with respect to the criterion resulting in a judgment matrix from which the vector of priorities is obtained using the principal eigenvector of the judgment matrix.

4.1. *Assembling the data*

The decision-making network of criteria and alternative outcomes is constructed as shown in Fig. 2. The object is to predict which team will win in their next matchup in a playoff game by incorporating expert judgment with historical data about the two teams. Table 3 shows the seven clusters of key factors, that are chosen using statistical methods as explained earlier. The factors are organized into clusters: offense, defense, positives, negatives, coach (teams' tactics), players' state, and other factors (e.g., season type, Home or Road game, trend, i.e., lucky streak). In the ANP network, factors in clusters are connected. For example, in the offense cluster, the Corsi (-) factor is affected by the tactics of the coach (in another cluster) and the physical and mental readiness of players (in another cluster) and the missed shots for (-) and blocked shots for (-) within the offense cluster. Another example, one of the offense factors is Corsi = shots on goal + missed shots + blocked shots. What's more, the nodes in the defense cluster affect other nodes inside the same cluster, as do the coach and players' state factors. There are similar relationships among other clusters.

We show here an example: the Chicago Blackhawks (CHI) and Tampa Bay Lightning (TBL) were compared for which is better regarding the physical factor. The "Coach Tactics" and "Players' State" nodes (physical and mental) are linked to the two teams and also to the data metrics. Thus, the data factors have one set of priorities resulting from comparisons of their importance with respect to the goal, and two more sets of priorities when they are compared with respect to Team A and Team B. In this feedback type of comparison, the question to be asked is: for the CHI, "Are they better at coach (tactics) or at Players state?". This is a subjective judgment that takes into account the current state of the team and how well the tactics are going lately. It results in two more sets of priorities for the two factors, one for CBH and another for TBL. Similar judgments are made throughout the network where the connections are. The limiting supermatrix of the ANP combines all the priorities in the network, and synthesizes them to give the outcome in the form of priorities for CHI and TBL. For instance, the synthesis might be 0.68 for Chicago

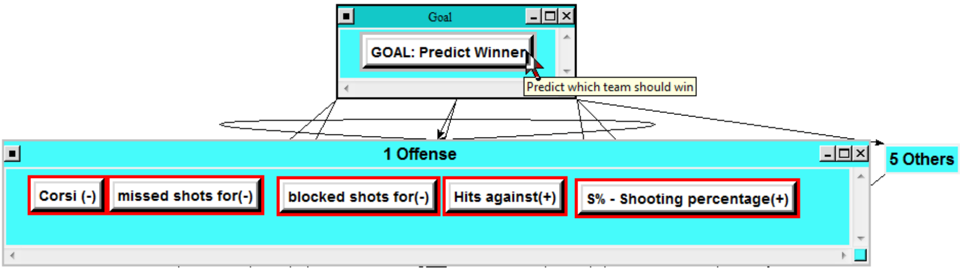

Fig. 3. Offense factors cluster expanded (factors linked from Goal).

and 0.32 for Tampa Bay. Thus the likelihood is that Chicago will win and that would be the prediction.

In Fig. 3, the goal is linked to the Offense cluster, which must be pairwise compared with respect to the goal for importance. In this particular set of comparisons, the judgments were made by experts in hockey games who arrived at their consensus judgments by examining data about the importance of these factors and interpreting it in the form of an AHP/ANP fundamental scale judgment (numbers from 1 to 9). The data about the factors was transformed into correlation coefficients derived from the aggregate performance of all the teams in the regular season. We detail it next.

4.2. *Judgment process for both the data and judgment factors*

In order to obtain judgments from ice hockey experts about games and teams based on their knowledge and experience, which is mostly not easy to obtain, we used a questionnaire format for the experts to fill out. In order to obtain judgments from ice hockey experts about games and teams based on their knowledge and experience, which is mostly not easy to obtain, we used a questionnaire format for the experts to fill out. Four experts are chosen to respond to the questionnaires. Two of them have studied and predicted ice hockey game for more than 10 years in Pittsburgh and the other two are ice hockey amateurs in Washington, D.C. with a great deal of expertise about the teams and sports. All the four respondents are very knowledgeable on NHL teams and ice hockey games. Namely, the experts participated in this research are knowledgeable and experienced.

The pairwise comparisons of relative importance of factors in all seven clusters are made before the playoff games start. The prediction of all games in the playoff is conducted before every game. As shown in Fig. 4, we ask experts to fill out questionnaires in order to elicit their subjective judgments combined with information from the correlation coefficients.

For example, blocked shots for -0.136 was considered moderately more important than Corsi for -0.076. A correlation coefficient of zero means not related. It is more important for a team to block a lot of shots than to receive a low Corsi score. Figure 4 exhibits the merged expert judgments with data.

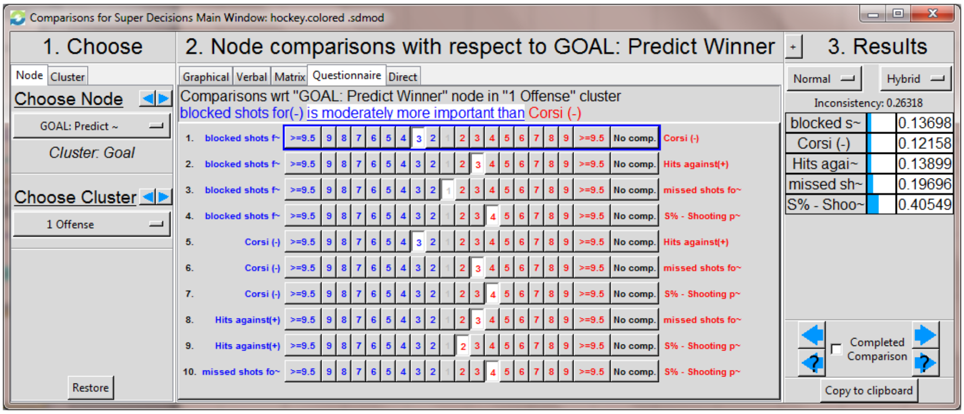

Fig. 4. Entering the judgments merged expert judgment with data.

As four experts are asked about the potential performance of each team involved in the playoff games, we need to integrate the judgments of all experts. In this study, we take experts as equal to derive group judgment through geometric mean; any inconsistency has been improved by consulting with the group when it occurs.

In this way, judgments were entered for all elements with respect to the goal. This resulted in one set of priorities for the criteria. The next step in our ANP model is to link all the criteria to the alternative teams (Chicago and Tampa Bay) and enter judgments based on comparing their statistical data (see Table 7). Saving percentages ALL vary in a narrow range (89.9% and 94.5%), so it takes the judgment of an expert, by looking at the data, to decide that Chicago (at 92.5%) is equally and moderately better than TB (at 91.1%).

The ice hockey commentators and pundits said Chicago was going to exert more physical pressure in this game. Information like this from various sources was considered and used to form the judgment inputs. The experts' judgments relied on data of past performance. Again we take the match between Chicago (CHI) and Tampa

Table 7. Match up data between CHI and TB in the 2014–2015 season.

	CHI	T.B		CHI	T.B
GF	289	324	BSF	1858	1511
GA	246	268	BSA	1401	1481
G+/-	43	56	SF%	51.7	51.2
Save%	92.5	91.1	SF	3500	3159
CF%	53.5	50.6	SA	3266	3016
CF%	53.5	50.6	Shoot%	8.3	10.3
CF	6597	5833	FO%	52.3	49.3
CA	5744	5685	HIT	2030	2594
MSF	1239	1163	HIT-	3293	2593
MSA	1077	1188	PenD	56	26

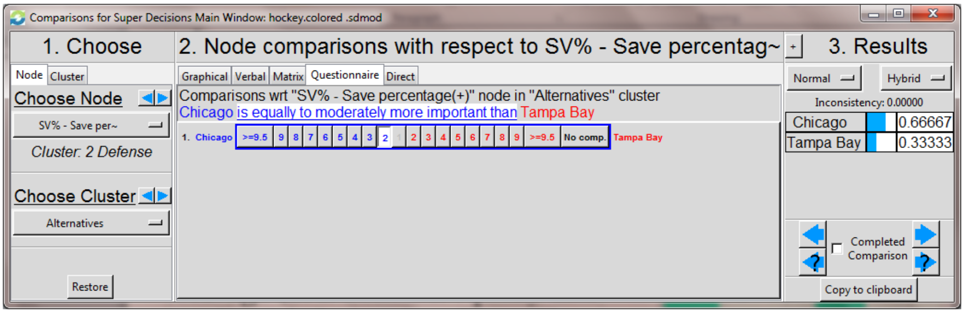

Fig. 5. Entering the merged expert judgments about "save percentage".

Bay (TBL) as an example to demonstrate the judgment process. First, there are some factors that are certain such as Home/Road game, Season type, for which experts can form judgments deterministically. Second, for other performance indicators, experts can make the judgment according to historical performance data (see Table 7), which is helpful for the experts to determine through this factor which team is better. In offence, the CHI has moderate advantage over TBL. While in defense, CHI is much stronger than TBL. Therefore, we have the judgment that TBL is equally to moderately more preferred (with the value 2 shown in Fig. 5) to CHI with respect to blocked shots; CHI is moderately more important (with the value 3 shown in Fig. 6) than TBL with respect to Corsi; and TBL is moderately more important (with the value of 3) than CHI with respect to Hits. In addition, there are no data to help the experts to make judgment about other indicators, such as trend to win/lose, mental state, and physical state, so they must use subjective judgment according to their best understanding of the game, of the teams, and of the current situation. As for the status of the teams, TBL can play more quickly, while CHI is slow according to historical performance. However, the players and teamwork of CHI are all very

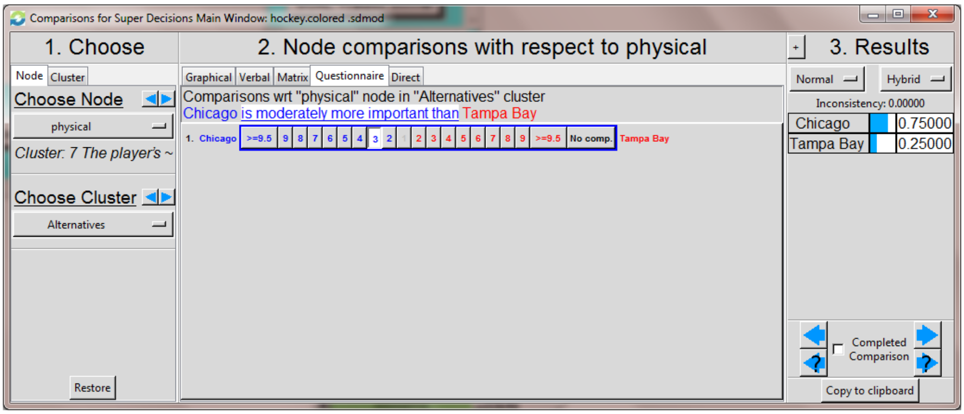

Fig. 6. Entering the merged expert judgment with data.

powerful particularly with groups 3 and 4, and also in the third period of the match when they are often able to reverse the outcome of the game. In the regular season games, the two teams played against each other twice, each winning one game. In the second game, without star player Patrick Kane CHI lost 0:4. Similarly, the judgments with respect to all the other elements were made. In the experiment, input judgments of the experts' consistency were all less than 0.1 which means the judgments were very reliable.

4.3. *Formation of the super matrix*

Assume that there are $p_1, \dots p_c, \dots p_n (c = 1, 2, \dots, n)$ criteria in the control level of the ANP model. We have a system of N clusters or components in network layer, whereby the elements in each component interact or have an impact on or are themselves influenced by some or all of the elements of that component or of another component with respect to a property governing the interactions of the entire system, such as energy or capital or political influence. Assume that component h, denoted by C_h, $h = 1, \dots, N$, has n_h elements, that we denote as $e_{h1}, e_{h2}, \dots, e_{hn_h}$. A priority vector derived from paired comparisons in the usual way represents the impact of a given set of elements in a component on another element in the system. When an element has no influence on another element, its influence priority is zero because there is no link to it.

The priority vectors derived from pairwise comparison matrices are each entered as a part of some column of a supermatrix. The supermatrix represents the influence

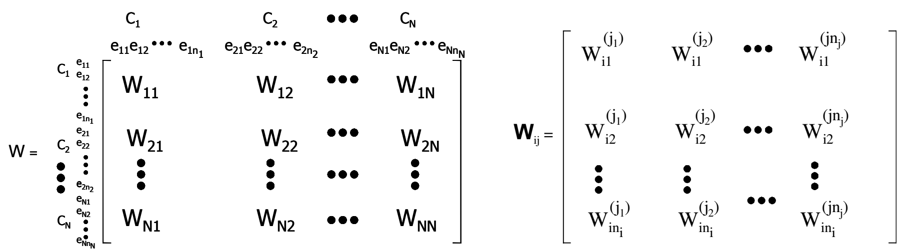

Fig. 7. The supermatrix of a network and detail of a matrix in it.

Table 8. Example of pairwise comparing offense elements for importance with respect to tactics.

	Blocked shots for(−)	Corsi(−)	Hits against(+)	Missed shots for(−)	S%(+)	Priorities
Blocked shots for(−)	1	1/5	1/3	1	1/9	0.0449
Corsi(−)	5	1	1/3	4	1/8	0.1222
Hits against(+)	3	3	1	5	1/5	0.1906
Missed shots for(−)	1	1/4	1/5	1	1/7	0.0448
S %(+)	9	8	5	7	1	0.5973

Note: Inconsistency index = 0.09710.

Table 9. Limit matrix.

	Blocked shots for (−)	Corsi (−)	Hits against(+)	Missed shots for (−)	S% (+)	Blocked shots against (+)	Corsi against (+)	Hits (−)	Missed shots against (+)	SV% (+)	Take away (+)
Blocked shots for(−)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Corsi (−)	0.4545	0.0000	0.0000	0.4545	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Hits against(+)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Missed shots for(−)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
S% (+)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Blocked shots against(+)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Corsi against(+)	0.0000	0.0000	0.0000	0.0000	0.0000	0.4545	0.0000	0.0000	0.4545	0.0000	0.0000
Hits(−)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Missed shots against(+)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
SV% (+)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Take away(+)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
FO% (+)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Power play opportunity(+)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Give away(−)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Penalty times(−)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Home/Road	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Season type	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Trend to win/lose	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Tactics	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Mentality	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Physical	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
A	0.3712	0.7500	0.6667	0.3712	0.6667	0.3636	0.6667	0.2500	0.3636	0.6667	0.3333
B	0.1742	0.2500	0.3333	0.1742	0.3333	0.1818	0.3333	0.7500	0.1818	0.3333	0.6667

Table 9. (*Continued*)

	FO% (+)	Power play (+)	Give away(−)	Penalty times(−)	Home/Road game	Season type	Trend to win/lose	Tactics	Mentality	Physical
Blocked shots for(−)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0070	0.0254	0.0183
Corsi (−)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0307	0.0576	0.0550
Hits against(+)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0297	0.0453	0.0473
Missed shots for(−)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0070	0.0181	0.0092
S %(+)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0930	0.0990	0.1062
Blocked shots against(+)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0322	0.0176	0.0173
Corsi against(+)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0494	0.0292	0.0251
Hits(−)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0056	0.0028	0.0032
Missed shots against(+)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0111	0.0118	0.0068
SV% - Save percentage(+)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1048	0.0476	0.0495
Take away(+)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0217	0.0082	0.0127
FO% (+)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0129	0.0172	0.0175
Power play opportunity(+)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0904	0.1201	0.1224
Give away(−)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0017	0.0014	0.0018
Penalty times(−)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0116	0.0110	0.0108
Home/Road game	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Season type	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Trend to win/lose	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Tactics	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Mentality	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Physical	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
A	0.7500	0.5000	0.5000	0.6667	0.2500	0.7500	0.6667	0.3056	0.3097	0.3127
B	0.2500	0.5000	0.5000	0.3333	0.7500	0.2500	0.3333	0.1859	0.1781	0.1843

priority of an element on the left of the matrix on an element at the top of the matrix. A supermatrix along with an example of one of its general entry i, j block are shown in Fig. 7. The component C_i alongside the supermatrix includes all the priority vectors derived for nodes that are "parent" nodes in the C_i cluster.

As an example for the pairwise comparisons here, tactics is one of the key factors that can influence the offence, defense, the negative and positive aspects of teams and games. Choose the Offence cluster, with respect to tactics, to take a pairwise comparison of the indicators. The comparisons are shown in Table 8. The priority column is in the unweighted supermatrix where the row is elements in Offence cluster and the head of the list is tactics which indicates significance of the influence with respect to tactics.

In the same way, input judgment of other experts' consistency test results are all less than 0.1 in which case the comparison matrix is acceptable and the unweighted supermatrix is formed.

Under the criteria p_c, make a comparison of the relative importance of the clusters $C_1, \ldots, C_N$, we will get the weighting matrix A of clusters. The cluster matrix is shown in Table 9. Then the weighted supermatrix is $\bar{W} = (\bar{W}_{ij})$ where $\bar{W}_{ij} = a_{ij} W_{ij}$, $i = 1, \ldots, N$, $j = 1, \ldots, N$.

The final priority of the elements is derived from raising the weighted supermatrix to powers until it stabilizes at the limiting supermatrix. In this case until each column is the same and identical. Solving the supermatrix is a very cumbersome and difficult process. The solutions in this paper were calculated by the Super Decisions (SD) Software. Table 8 gives a pairwise comparison judgment matrix along with its calculated priority vector. The limit matrix of our model is shown in Table 9.

4.4. *Example of predicting the outcome of a match*

In the case under study, the alternatives are two teams who are about to face each other in an upcoming match. The object is to pick the winning team. Table 10 shows

Table 10. The result of the example of the game between CHI and TBL.

Name	Normalized by cluster	Limiting
Blocked shots for(−)	0.07634	0.01186
Corsi (−)	0.22482	0.034926
Hits against(+)	0.18672	0.029007
Missed shots for(−)	0.07981	0.012398
Shooting percentage(+)	0.43232	0.067161
Blocked shots against(+)	0.14539	0.019623
Corsi against(+)	0.22575	0.030468
Hits(−)	0.03345	0.004515
Missed shots against(+)	0.05373	0.007252
Save percentage(+)	0.43743	0.059038
Take away(+)	0.10425	0.01407

Table 10. (*Continued*)

Name	Normalized by cluster	Limiting
Faceoff win percentage(+)	0.13868	0.009934
Power play opportunity(+)	0.86132	0.061698
Give away(−)	0.15759	0.001973
Penalty times(−)	0.84241	0.010547
Home/Road game	0.23639	0.00211
Season type	0.0819	0.000731
Trend to win/lose	0.68172	0.006085
Tactics	1	0.117642
Mentality	0.66667	0.091008
Physical	0.33333	0.045504
CHI	0.62822	0.227697
TB	0.37178	0.134753

Table 11. Comparison of prediction accuracy of sports game.

Paper work	Sports	Data	Method	Accuracy
Weissbock and Inkpen[6]	Ice hockey	720 NHL games in the 2012–2013 NHL shortened season	Machine learning	60.25%
Morgan *et al.*[9]	Hockey	Eight elite female hockey players in two sessions	Machine learning	64.3%
Huang and Chang[25]	Soccer	2006 World Cup	Neural networks	76.9%
Kahn, 2003	Football	NFL football game for the 2003	Neural networks	75%
Loeffelholz *et al.*[13]	Basketball	620 NBA games	Neural networks	74.33%
Miljkovic *et al.* 2010	Basketball	NBA	Naive Bayes Multivariate linear regression	67%
Yang and Lu, 2012	Basketball	NBA	SVM	86.75%
Weissbock *et al.*[2]	Ice Hockey	2012–2013 season for a total of 517 games between 16 February and 28 April 2013	SVM	59.8%
Our paper	Ice hockey	1,230 regulation games in NHL 2014–2015 season for predicting 89 post season games	Data and judgment	77.5%

the overall result. Under the "Normalized by cluster" column, we find CHI will win over Tampa Bay with a 63% advantage. Thus, we predict that CHI will win. After the actual match which score was Chicago 2, and Tampa Bay 1 (which is fairly close to 0.66 to 0.33 of our prediction). After the actual match, we collected information about the game, shown in Table 12, and we compared it with prediction which we had made prior to the game.

Table 12. The information of practical match of the CHI and NBL.

Date	Visitor	Home	R/H			Prediction	Dec	Sv%/sv%-a		CF/CA		MSF/MSA		BSF/BSA		SF/SA		Sh%	FO%	HIT/HIT-				PenD
2015/6/3	CHI	T.B	R				W	>1		>1		>1		<1		>1		9.5	>50%	<1				>0
				GF	GA	outcome	Dec	Sv%	sv%-a	CF	CA	MSF	MSA	BSF	BSA	SF	SA	Sh%	FO%	HIT	HIT-	PN	PN-	PenD
				2	1		W	95.7	90.5	52	45	16	11	15	11	21	23	9.5	47.4	21	29	2	3	1

Table 13. The prediction of NHL in 2014–2015 season playoff games.

Date	Visitor	Home	R/H	GF	GA	Prediction	Dec	Sv%	sv%-a	CF	CA	MSF	MSA	BSF	BSA	SF	SA	Shoot%	FO%	HIT	HIT-	PN	PN-	PenD
2015/4/15	NYI	WSH	R	4	1	W	W	96	85.2	65	55	21	11	17	19	27	25	14.8	37.1	36	46	2	3	1
2015/4/15	CHI	NSH	R	4	3	W	W	94.4	90.5	93	84	19	14	32	16	42	54	9.5	38.6	28	39	6	6	0
2015/4/15	OTT	MTL	R	3	4	W	L	89.7	90.9	70	63	16	10	21	14	33	39	9.1	49.4	44	42	3	4	1
2015/4/15	CGY	VAN	R	2	1	W	W	96.7	93.3	54	61	11	11	13	20	30	30	6.7	45.8	25	22	3	1	−2
2015/4/16	PIT	NYR	R	1	2	L	L	94.7	96.0	51	55	4	6	22	11	25	38	4	57.6	32	32	5	1	−4
2015/4/16	MIN	STL	R	4	2	W	W	90.5	86.2	49	58	6	17	14	20	29	21	13.8	41.9	25	22	2	5	3
2015/4/16	WPG	ANA	R	2	4	L	L	87.9	92.6	42	67	9	20	6	14	27	33	7.4	37.5	37	46	4	3	−1
2015/4/16	DET	T.B	R	3	2	L	W	95.7	78.6	28	68	6	9	8	13	14	46	21.4	41.3	25	38	7	4	−3
2015/4/17	CHI	NSH	R	2	6	W	L	82.9	92.3	65	59	16	8	23	16	26	35	7.7	51.2	25	19	6	4	−2
2015/4/17	CGY	VAN	R	1	4	L	L	87.5	95.7	46	56	11	9	12	15	23	32	4.3	50	19	30	9	4	−5
2015/4/17	OTT	MTL	R	2	3	L	L	92.9	93.5	57	80	8	14	18	24	31	42	6.5	47	53	41	6	3	−3
2015/4/17	NYI	WSH	R	3	4	L	L	88.6	85.7	51	82	19	20	11	27	21	35	14.3	49.1	38	59	1	1	0
2015/4/18	DET	T.B	R	1	5	L	L	83.3	95.8	58	48	13	8	21	10	24	30	4.2	55.8	25	34	6	6	0
2015/4/18	MIN	STL	R	1	4	L	L	85.2	96.0	47	50	10	9	12	14	25	27	4	42.3	36	36	3	1	−2
2015/4/18	WPG	ANA	R	1	2	L	L	94.9	96.6	56	62	12	15	15	8	29	39	3.4	47.5	33	48	5	4	−1
2015/4/18	PIT	NYR	R	4	3	L	W	88.5	81.8	37	62	8	16	7	20	22	26	18.2	50	21	25	7	4	−3
2015/4/19	MTL	OTT	R	2	1	L	W	97.1	95.9	89	65	15	11	25	20	49	34	4.1	44.8	36	61	3	7	4
2015/4/19	VAN	CGY	R	2	4	L	L	85.2	92.0	65	56	10	10	30	19	25	27	8	50	18	32	9	7	−2
2015/4/19	NSH	CHI	R	2	4	L	L	86.7	94.6	62	59	8	9	17	20	37	30	5.4	42.4	53	37	3	3	0
2015/4/19	WSH	NYI	R	1	2	L	L	95.2	96.0	57	76	8	10	24	24	25	42	4	50.8	43	44	3	3	0
2015/4/20	NYR	PIT	R	2	1	L	W	95.8	92.3	50	55	10	14	14	17	26	24	7.7	45.9	37	43	4	5	1

Table 13. (*Continued*)

Date	Visitor	Home	R/H	GF	GA	Prediction	Dec	Sv%	sv%-a	CF	CA	MSF	MSA	BSF	BSA	SF	SA	Shoot%	FO%	HIT	HIT-	PN	PN-	PenD
2015/4/20	ANA	WPG	R	5	4	W	W	88.6	83.3	56	62	9	12	17	15	30	35	16.7	48.8	44	61	4	4	0
2015/4/20	STL	MIN	R	0	3	L	L	87.5	100.0	47	55	12	12	18	19	17	24	0	50.8	26	28	4	0	−4
2015/4/21	VAN	CGY	R	1	3	L	L	86.4	96.6	70	38	12	7	29	9	29	22	3.4	46.8	18	29	4	3	−1
2015/4/21	T.B	DET	R	0	3	L	L	85.7	100.0	41	34	9	8	10	5	22	21	0	50	26	48	5	7	2
2015/4/21	WSH	NYI	R	2	1	L	W	97.3	93.3	66	78	12	20	24	21	30	37	6.7	50.8	38	50	4	1	−3
2015/4/21	NSH	CHI	R	2	3	L	L	93.8	96.2	78	105	7	18	19	39	52	48	3.8	46.6	66	50	5	4	−1
2015/4/22	STL	MIN	R	6	1	W	W	94.4	76.9	57	43	22	10	9	15	26	18	23.1	50	16	26	1	2	1
2015/4/22	NYR	PIT	R	2	1	L	W	95.7	91.7	50	50	9	8	17	19	24	23	8.3	42.9	42	28	3	3	0
2015/4/22	ANA	WPG	R	5	2	W	W	92.6	85.7	54	54	10	12	9	15	35	27	14.3	53.5	35	45	3	1	−2
2015/4/22	MTL	OTT	R	0	1	W	L	96.9	100.0	43	66	4	17	11	17	28	32	0	45.1	29	34	2	3	1
2015/4/23	T.B	DET	R	3	2	W	W	91.7	89.7	46	44	7	8	10	12	29	24	10.3	55.4	34	30	6	6	0
2015/4/23	CHI	NSH	R	2	5	W	L	82.8	93.3	57	60	12	17	15	14	30	29	6.7	46.9	16	36	7	5	−2
2015/4/23	NYI	WSH	R	1	5	L	L	87.8	95.7	60	67	19	10	18	16	23	41	4.3	36.7	49	49	6	3	−3
2015/4/23	CGY	VAN	R	1	2	L	L	95.3	95.2	50	66	12	7	17	16	21	43	4.8	50.8	25	15	2	3	1
2015/4/24	PIT	NYR	R	1	2	W	L	94.4	97.4	74	65	17	16	19	13	38	36	2.6	52.9	28	36	3	3	0
2015/4/24	OTT	MTL	R	5	1	W	W	97.8	80.0	49	82	5	10	19	26	25	46	20	45.6	50	39	5	6	1
2015/4/24	MIN	STL	R	4	1	W	W	97.3	78.9	39	72	7	17	13	18	19	37	21.1	46.8	19	22	3	3	0
2015/4/25	VAN	CGY	R	4	7	L	L	78.8	81.8	58	55	13	11	23	11	22	33	18.2	50.7	15	32	3	2	−1
2015/4/25	WSH	NYI	R	1	3	L	L	92.1	97.4	65	66	12	10	14	18	39	38	2.6	54.8	32	46	7	10	3
2015/4/25	NSH	CHI	R	3	4	L	L	87.5	88.0	42	55	7	8	10	15	25	32	12	50	53	36	3	3	0
2015/4/25	DET	T.B	R	4	0	L	W	100	86.7	46	59	6	17	10	14	30	28	13.3	35.9	32	26	4	4	0
2015/4/26	STL	MIN	R	1	4	L	L	81	96.8	65	38	11	11	23	6	31	21	3.2	47.2	12	15	2	3	1
2015/4/26	MTL	OTT	R	2	0	W	W	100	90.0	44	70	15	12	9	15	20	43	10	58.5	23	45	4	1	−3
2015/4/27	T.B	DET	R	5	2	W	W	91.7	82.1	43	53	7	16	8	13	28	24	17.9	55.2	27	29	11	6	−5
2015/4/27	NYI	WSH	R	1	2	L	L	92.3	90.9	47	60	18	14	18	20	11	26	9.1	39.6	53	46	0	1	1

Table 13. (*Continued*)

Date	Visitor	Home	R/H	GF	GA	Prediction	Dec	Sv%	sv%-a	CF	CA	MSF	MSA	BSF	BSA	SF	SA	Shoot%	FO%	HIT	HIT-	PN	PN-	PenD
2015/4/29	DET	T.B	R	0	2	L	L	88.2	100.0	56	33	9	8	16	8	31	17	0	50.9	31	22	5	4	−1
2015/4/30	WSH	NYR	R	2	1	W	W	96.9	93.1	60	65	16	14	15	19	29	32	6.9	44.3	32	34	3	3	0
2015/4/30	CGY	ANA	R	1	6	L	L	82.9	95.8	57	64	14	15	19	14	24	35	4.2	48.4	22	27	6	5	−1
2015/5/1	MIN	CHI	R	3	4	K	L	88.6	90.9	52	55	7	8	12	12	33	35	9.1	44.1	36	34	1	3	2
2015/5/1	T.B	MTL	R	2	1	W	W	97.7	94.3	77	85	8	14	34	27	35	44	5.7	38.2	32	43	3	4	1
2015/5/2	WSH	NYR	R	2	3	L	L	91.4	93.8	59	63	15	13	12	15	32	35	6.2	51.7	38	30	4	1	−3
2015/5/3	CGY	ANA	R	0	3	L	L	91.2	100.0	60	63	10	11	20	18	30	34	0	46	35	39	5	4	−1
2015/5/3	MIN	CHI	R	1	4	L	L	87.1	96.8	51	63	11	14	9	18	31	31	3.2	47.5	42	39	2	2	0
2015/5/3	T.B	MTL	R	6	2	W	W	93.1	75.0	35	54	1	9	10	16	24	29	25	45	21	28	5	13	8
2015/5/4	NYR	WSH	R	0	1	L	L	95.5	100.0	69	49	12	15	27	12	30	22	0	31	31	39	2	2	0
2015/5/5	ANA	CGY	R	3	4	W	L	81	85.7	45	57	15	18	9	18	21	21	14.3	67.8	21	26	7	3	−4
2015/5/5	CHI	MIN	R	1	0	W	W	100	95.5	41	65	9	16	10	19	22	30	4.5	62.7	11	20	3	1	−2
2015/5/6	MTL	T.B	R	1	2	L	L	89.5	96.8	69	41	21	12	17	10	31	19	3.2	51.7	24	28	3	2	−1
2015/5/6	NYR	WSH	R	1	2	L	L	93.3	96.6	66	49	12	12	25	7	29	30	3.4	43.3	31	37	5	7	2
2015/5/7	MTL	T.B	R	6	2	L	W	91.7	85.0	56	41	6	6	10	11	40	24	15	51.5	29	30	4	4	0
2015/5/7	CHI	MIN	R	4	3	W	W	91.9	84.0	53	64	14	13	14	14	25	37	16	50	6	24	3	2	−1
2015/5/8	ANA	CGY	R	4	2	W	W	92.6	86.2	67	52	15	11	23	14	29	27	13.8	49.3	25	27	3	4	1
2015/5/8	WSH	NYR	R	1	2	W	L	95.3	96.6	56	66	6	10	21	13	29	43	3.4	52.3	24	32	2	2	0
2015/5/9	T.B	MTL	R	1	2	W	L	93.1	96.0	54	62	10	16	19	17	25	29	4	47.4	44	30	5	2	−3
2015/5/10	NYR	WSH	R	4	3	L	W	93.3	85.7	55	96	12	17	15	34	28	45	14.3	45.9	30	36	4	3	−1
2015/5/10	CGY	ANA	R	2	3	L	L	93.6	89.5	40	90	12	16	9	27	19	47	10.5	37.7	31	33	6	5	−1
2015/5/12	MTL	T.B	R	1	4	L	L	85.7	94.7	40	47	10	7	11	12	19	28	5.3	41.9	22	33	2	2	0
2015/5/13	WSH	NYR	R	1	2	L	L	94.9	97.2	68	74	13	14	19	21	36	39	2.8	59.5	26	32	4	3	−1
2015/5/16	T.B	NYR	R	1	2	L	L	93.3	95.8	50	61	11	14	15	17	24	30	4.2	56.6	18	30	2	4	2
2015/5/17	CHI	ANA	R	1	4	L	L	85.2	97.0	70	53	15	17	22	9	33	27	3	48.1	34	44	1	3	2

Table 13. (*Continued*)

Date	Visitor	Home	R/H	GF	GA	Prediction	Dec	Sv%	sv%-a	CF	CA	MSF	MSA	BSF	BSA	SF	SA	Shoot%	FO%	HIT	HIT-	PN	PN-	PenD
2015/5/18	T.B	NYR	R	6	2	W	W	94.6	76.9	41	63	5	9	10	17	26	37	23.1	45.6	26	28	6	7	1
2015/5/19	CHI	ANA	R	3	2	W	W	96.8	94.6	115	129	24	38	35	29	56	62	5.4	47.3	45	71	5	5	0
2015/5/20	NYR	T.B	R	5	6	L	L	85	82.1	58	64	17	10	13	14	28	40	17.9	45.8	33	41	3	5	2
2015/5/21	ANA	CHI	R	2	1	L	W	96.4	92.6	46	67	10	12	9	27	27	28	7.4	48.5	45	27	5	2	−3
2015/5/22	NYR	T.B	R	5	1	W	W	97.4	79.2	42	61	9	13	9	9	24	39	20.8	43.6	30	32	9	10	1
2015/5/23	ANA	CHI	R	4	5	L	L	87.5	92.2	78	92	7	18	20	34	51	40	7.8	45.5	60	52	5	3	−2
2015/5/24	T.B	NYR	R	2	0	W	W	100	90.9	42	58	9	8	11	24	22	26	9.1	61.5	29	29	5	4	−1
2015/5/25	CHI	ANA	R	4	5	L	L	82.1	85.7	66	54	17	15	21	11	28	28	14.3	40	23	41	2	2	0
2015/5/26	NYR	T.B	R	7	3	L	W	92.3	79.4	55	65	9	9	12	17	34	39	20.6	53	27	22	4	4	0
2015/5/27	ANA	CHI	R	2	5	L	L	78.3	93.8	57	54	6	8	19	23	32	23	6.2	34	43	38	3	3	0
2015/5/29	T.B	NYR	R	2	0	W	W	100	92.0	44	53	7	11	12	20	25	22	8	37.5	25	29	2	0	−2
2015/5/30	CHI	ANA	R	5	3	W	W	92.1	80.8	51	68	11	15	14	15	26	38	19.2	50	15	37	2	4	2
2015/6/3	CHI	T.B	R	2	1	W	W	95.7	90.5	52	45	16	11	15	11	21	23	9.5	47.4	21	29	2	3	1
2015/6/6	CHI	T.B	R	3	4	L	L	83.3	89.7	49	48	11	12	9	12	29	24	10.3	64.8	28	33	3	3	0
2015/6/8	T.B	CHI	R	3	2	L	W	94.7	90.6	51	67	5	10	14	19	32	38	9.4	41.8	46	27	3	3	0
2015/6/10	T.B	CHI	R	1	2	L	L	89.5	96.0	61	45	13	9	23	17	25	19	4	34.5	46	34	3	4	1
2015/6/13	CHI	T.B	R	2	1	W	W	96.9	93.1	56	61	11	11	16	18	29	32	6.9	58.1	15	37	1	2	1
2015/6/15	T.B	CHI	R	0	2	L	L	93.8	100.0	60	58	10	14	25	12	25	32	0	32.3	56	32	3	1	−2

5. Results and Discussion

Using data analysis and judgment models, we predicted the outcomes of all the 2014–2015 season playoff games. The prediction results are shown in Table 13. There are 89 games, of which we predicted in advance with 69 correct game outcomes, giving us an accuracy rate of 77.5%. Compared with the results found by the University of Ottawa, ours appeared to achieve better accuracy, i.e., better than the upper bound of 62% they stated.

When analyzing data of the NHL 2014–2015 season, factor analysis approach could be used to show that the star players, which is primarily what people are concerned with, are actually not determinant attributes in the game. We did this, but the explanation is long and we do not include it in this paper.

In terms of players' state in the game, the mentality of players is more important than the physical state. The mentality, we proposed represents psychological factors, intelligence factors and sentiment as well. The physical power of players, if they are not totally disabled, are subordinate to their mental state. Psychological factors play an important role in the game as favorable mentality will improve attention, confidence and the capacity to endure strain. That's why adjusting the players' mentality before the game and consciously cultivating their mental attitude during training exercises should be taken seriously. The result concludes that the state of players and teams should also be taken into account in predicting sports outcomes. The SVM result was important to demonstrate that the factors we included in the paper can achieve a high degree of accuracy in prediction.

For more comparisons of how our results stand compared to others, we refer to previous literature about prediction in sports (see Table 11). Basketball, football and soccer can attain a higher accuracy in prediction because they have many more events than does ice hockey, so it can lead to higher accuracy in prediction. Though SVM achieved a very high accuracy of 86.75% in predicting basketball, the accuracy lowered to 59.8% when predicting hockey. We can see that ML techniques are limited to data and depend on the sensitivity of the input information. What is more, in the previous ice hockey predictions, the data are from published reports. Compare that to the judgment process, this paper directly interpreted primary data from experts.

6. Conclusion

In this paper, we collected data of 1230 NHL regular season games in 2014–2015 from diverse sources on players, teams and game performances. We filter and reduce the 54 metrics to 17 factors that were important for estimating game performance through rank-sum tests. These factors then become the tangible criteria in the proposed expert system. Next, we confirm that the chosen indicators are indeed significant by applying SVM to validate the regular season games for which the outcomes were known. Integrated with tangible performance metrics derived in Sec. 3, we incorporate intangible criteria (i.e., coach tactics, mental and physical states) into the expert system. In this way, the ANP model incorporates both data and judgments.

We use the proposed ANP model to predict each of the 89 post-season NHL games in 2014–2015 season. The results are quite accurate. Specifically, we correctly predict 77.5% of the matches before they took place. We hope to test the methodology further in the coming seasons. Compared with previous methods used in sports prediction, the model we proposed has the following advantages:

(1) We are the first to predict hockey games using a combination of both tangible data and judgments on intangibles. Prior to this study, there were predictions-based solely on tangible data, or predictions by commentators that relied solely on intangibles to speculate which team is more likely to win. Our research is the first to offer an effective and efficient model that successfully combines both tangible and intangible information. This study has the potential to make full use of the knowledge and experience of experts, while also incorporating ML and surpassing its forecasting accuracy. In addition, the expert system framework which makes use of factor analysis, SVM, and the ANP, can be employed in other settings, where intangible and tangible indicators are earnestly needed to produce better estimates.
(2) The factors used in our prediction have been validated for their usefulness in producing accurate outcomes. Using rank-sum tests, we showed that the criteria used in the ANP model significantly influence the game results. Furthermore, we use SVM to confirm that the factors adopted can derive high accurate prediction. Finally, through correlation coefficients between the set of tangible factors and win/loss outcome, we know the relative importance of each factor. Compared with other ML methods which depend on the text analysis of collected sports comments, the proposed expert system based on both data and judgment seems more reliable, as forecasting accuracy has been greatly improved.
(3) The objective judgments in the model were evidence-based, in which the experts use the information from the numerical analysis to support their judgments (e.g., correlation). When the experts evaluate the two teams based on factors such as shooting percentage, shots, hits, giveaways, and takeaways, they are making more structured and fact-based assessment, which often results in higher accuracy. Along with subjective judgments such as commentators' views, interpretations of the players physical/mental status, and their general knowledge of the game. The expert system offers a more transparent and justifiable assessment framework and insights for sports performance prediction.

There are also limitations in this research that need to be improved in future work.

(1) To ensure minimum change in the teams and players between the regular season and the playoff season, we limited our NHL data to just one season (2014–2015). In subsequent study, more historical data will be taken into consideration.

(2) When looking for factors that influence game outcomes, we studied all games of all teams in NHL. Thus, the key factors identified are common factors for all teams. However, for a specific game, if we focus on the two competing teams' historical data, we may be able to identify distinctive factors uniquely suitable for such match and better able to predict the game outcome.

(3) The intangible judgments heavily relied on the experts' backgrounds, knowledge and experience. In the future, we can explore whether subjective judgments could be further improved with additional evidence-based historical data. In future research into sports game prediction, big data[24] may provide richer references for experts to make the judgments. We can also collect the judgments from experts then use group decision-making approaches to combine their judgments. By using group questionnaires and synthesizing the judgments from more experts with access to evidence-based data, we may be able to achieve ever-higher accuracy.

Acknowledgment

This research is supported by the National Natural Science Foundation of China (71572011), the fundamental Research Funds for the Central Universities (FRF-BR-15-011B).

References

1. N. Bohr, *Atomic Physics and the Description of Nature* (Cambridge University Press, England, 1934).
2. J. Weissbock, H. Viktor and D. Inkpen, Use of performance metrics to forecast success in the national hockey league, *Paper presented at the Workshop on Sports Data Mining at ECML/PKDD 2013* (2013).
3. T. L. Saaty and L. Zhang, (2016). The need for both Bayes' theorem and judgment in prediction (in process).
4. D. Voyer and E. F. Wright, Predictors of performance in the national hockey league, *Journal of Sport Behavior* **21**(4) (1998) 456.
5. J. Weissbock, *Forecasting Success in the National Hockey League Using In-Game Statistics and Textual Data* (University of Ottawa, Canada, 2014).
6. J. Weissbock and D. Inkpen, Combining textual pre-game reports and statistical data for predicting success in the national hockey league, *Lecture Notes in Computer Science* (2014) 251–262.
7. C. A. Depken, R. Sonora and D. P. Wilson, Performance under pressure: Preliminary evidence from the national hockey league. *International Journal of Sport Finance*, **7**(3) (2012) 213–231.
8. G. I. Kolev, G. Pina and F. Todeschini, Decision making and underperformance in competitive environments: Evidence from the national hockey league, *Kyklos* **68**(1) (2015) 65–80.
9. S. Morgan, M. D. Williams and C. Barnes, Applying decision tree induction for identification of important attributes in one-versus-one player interactions: A hockey exemplar, *Journal of Sports Science* **31**(10) (2013) 1031–1037.

10. H. Chen, P. Buntin Rinde, L. She, S. Sutjahjo, C. Sommer and D. Neely, Expert prediction, symbolic learning, and neural networks. An experiment on greyhound racing, *IEEE Expert* **9**(6) (1994) 21–27.
11. K. Huang and K. Chen, Multilayer perceptron for prediction of 2006 world cup football game, *Advances in Artificial Neural Systems*, **2011** (2011) 1–8.
12. K. Y. Huang and W. L. Chang, A neural network method for prediction of 2006 world cup football game, *In the 2010 International Joint Conference on Neural Networks (IJCNN)*, IEEE, 1–8 (2010).
13. B. Loeffelholz, E. Bednar and K. Bauer, Predicting NBA games using neural networks, *Journal of Quantitative Analysis in Sports* **5**(1) (2009) 7.
14. M. van Roon, *Predicting the Outcome of NBA Playoffs* (Vrije Universiteit, Amsterdam, 2012).
15. A. Zimmermann, S. Moorthy and Z. Shi, Predicting college basketball match outcomes using machine learning techniques: Some results and lessons learned, (2013), arXiv preprint: 1310.3607.
16. D. L. Feltz and C. D. Lirgg, Perceived team and player efficacy in hockey, *Journal of Applied Psychology* **83**(4) (1998) 557–564.
17. A. H. Perlini and T. R. Halverson, Emotional intelligence in the national hockey league, *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement* **38**(2) (2006) 109–119.
18. S. E. Buttrey, A. R. Washburn and W. L. Price, Estimating NHL scoring rates, *Journal of Quantitative Analysis in Sports* **7**(3) (2011).
19. B. Macdonald, An expected goals model for evaluating NHL teams and players, *Paper Presented at the MIT Sloan Sports Analytics Conf. 2012* (2012).
20. T. L. Saaty, *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation* (McGraw-Hill, US, 1980).
21. T. L. Saaty, *Decision Making with Dependence and Feedback: The Analytic Network Process: The Organization and Prioritization of Complexity* (RWS Publication, US, 1996).
22. T. L. Saaty, The analytic network process, In *Decision Making with the Analytic Network Process* (Springer, US, 2006), pp. 1–26.
23. T. L. Saaty, The analytic hierarchy and analytic network processes for the measurement of intangible criteria and for decision-making, *International Series in Operations Research & Management Science*, Vol. 78, (Springer, New York, 2005), pp. 345–405.
24. J. Fan, F. Han and H. Liu, Challenges of big data analysis, *National Science Review* **1**(2) (2014) 293–314.
25. K. Y. Huang and W. L. Chang, A neural network method for prediction of 2006 world cup football game in best-of-seven playoff series, *Paper presented at the 2010 Int Joint Conf. on Neural Networks* (IJCNN), (2010).
26. G. Pischedda, Predicting NHL match outcomes with ML models, *International Journal of Computer Applications* **101**(9) (2014) 15–22.
27. J. C. Platt, *Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines* (Microsoft Research, 1998).
28. T. L. Saaty, *Decision Making for Leaders* (Lifetime Learning Pubications, CA, 1982).
29. J. Kahn, *Neural Network Prediction of NFL Football Games* (World Wide Web electronic publication, 2003), pp. 9–15.
30. D. Miljkovic, L. Gajic, A. Kovacevic and Z. Konjovic, The use of data mining for basketball matches outcomes prediction, Paper presented at the *IEEE 2010 8th International Symposium on Intelligent Systems and Informatics (SISY)*, (2010).
31. J. B. Yang and C. Lu, Predicting NBA championship by learning from history data, Paper presented at the *Proceedings of Artificial Intelligence and Machine Learning for Engineering Design* (2012).