International Journal of Operations Research

TECHINAL NOTE

International Journal of Operations Research Vol. 5, No. 3, 187-192 (2008)

Clinching and Elimination of Playoff Berth in the NHL

Eddie Cheng [1,*] **and Daniel E. Steffy** [2]

[1]Department of Mathematics and Statistics, Oakland University, Rochester, Michigan, 48309

[2] H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia, 30332

Received May 2006; *Accepted February 2007*

Abstract**—**This paper considers the problem of determining as early as possible when teams in the NHL are mathematically qualified or eliminated from the playoffs. A mixed integer program formulation to determine mathematical qualification and elimination will be given via the rules outlined by the NHL. We also give a slightly simplified version of the MIP that is more computationally tractable and computational results for the 2003–2004 NHL season in which our method is able to detect qualification and elimination earlier than the rules currently used by the sports media.

Keywords**—**Integer program, Formulation, NHL, Applications

1. INTRODUCTION

The National Hockey League (NHL) is the organization of professional ice hockey teams in the United States and Canada. For the 2003-2004 season, it had 30 teams divided into two conferences (Eastern and Western) with 15 teams in each of them. Each conference is subdivided into three divisions with 5 teams in each divisions. During the regular season, each team plays 82 games (41 home games and 41 road games). Each team plays against a team from the other conference once or twice, against a team from the same conference but a different division three or four times, and against a team within the same division five or 6 times.

When two teams play against each other and the score is tied at the end of the game, they will play a 5-minute overtime. Hence a game between Team A and Team B has five possible outcomes: Team A beats Team B at the end of regulation, they tie at the end of regulation and Team A beats Team B in overtime, and they tie at the end of regulation and at the end of overtime, or Team B beats Team A regularly or in overtime. In the first case, Team A receives two points and Team B receives none. In the second case, Team A receives two points and Team B receives one point. In the third case, Team A and Team B each receives one point and the cases where Team B wins are similar. This award system is summarized in Table 1. (Starting in the 2005–2006 season, overtime tie games are eliminated, so (1 point, 1 point) distribution is no longer available. Our formulations in this paper are still valid as the new rule is a simplification of the old rule. In fact it will be seen later that tie games need not be considered with our method.)

At the end of the regular season, the playoffs begin. The team with the most points in each division is the division leader. Eight teams from each conference are eligible for the playoffs with the three division leaders automatically qualified for the playoff. The other five teams are chosen from the remaining teams with the most points. Of course, this method does not uniquely determine the eight teams as it is possible for two teams to have the same number of points. So a tie breaking rule is used[1]. If two (or more) teams have the same number of points, a tie is broken by considering the total number of wins each team has. If they are still tied, such a tie is broken by the number of points each team has earned against the other tied team or teams, then even further the number of wins against these teams. To avoid confusion the points referred to in these rules are the points scored via the outcome of each game, not to be confused with the actual goals scored during any of the games. If they are still tied, such a tie is broken by looking at the difference between goal scored and against among the tied teams. Of course it is possible that they are still tied, for example, every game during the regular season is tied after overtime with a score of 0 to 0. In this case, further tie breaking rules would be decided by the NHL.

Table 1. Points awarded to teams

	Team A	Team B
Team A wins Regular game	2	0
Team A wins in Overtime	2	1
Tie Game	1	1
Team B wins in Overtime	1	2
Team B wins Regular game	0	2

One interesting question that is closely followed by the media and fans is whether a team has clinched a playoff spot or has been eliminated from the playoff from the current record before the end of regular season.

* Corresponding author's email: echeng@oakland.edu

[1]Tie breaking rule can be obtained from www.espn.com or http://proicehockey.about.com/library/blstandings.htm

Table 2. Ranking and tie breaking rules

Initial scores	Points awarded based on outcome of games
Level 1 tie breaking	Wins
Level 2 tie breaking	Points earned against set of teams tied in Level 1 tie breaking
Level 3 tie breaking	Wins among set of teams tied in Level 2 tie breaking

We say a team is *mathematically qualified for the playoff with respect to* $\mathcal{R}$ where $\mathcal{R}$ is a set of rules if it will be qualified regardless of the outcome of the remaining games by applying the rules in $\mathcal{R}$, and a team is *mathematically eliminated from the playoff with respect to* $\mathcal{R}$ if it cannot make the playoff regardless of the outcome of the remaining games by applying the rules in $\mathcal{R}$. We note that $\mathcal{R}$ is important in the definition. For example, one may want to simplify the NHL regulation and assume there is no mathematical tie breaking rule (tied teams are drawn at random to fill the playoff spots), then a team is mathematically qualified for the playoff with respect to this rule if it is guaranteed that its total point is strictly greater than all but seven teams including the division leader. This condition is of course sufficient but too strong to be of any practical value. In fact, our experiment shows that this simplification is often worse than the simple rule used by the media. If more levels of tie breaking rule are included in $\mathcal{R}$, the sufficiency needed to declare a team is mathematically qualified or eliminated will be weaker, that is, one may discover the truth earlier. We let $\mathcal{R}_i$ denotes the rule of including tie breaking rule up to and including level *i* given in Table 2.

In this paper, we consider the formulation of the qualification and elimination problems with respect to $\mathcal{R}_1$ and $\mathcal{R}_3$. Our formulation is an integer linear programming formulation. In 2005 Ribeiro and Urrutia give a formulation for the Brazilian Soccer League which motivated our research. The Brazilian Soccer League playoff qualification rules are simpler than those of the NHL and their season involves fewer games. Although the rules for the NHL are more complicated, a mixed integer programming formulation for the qualification and elimination problems with respect to $\mathcal{R}_1$ as well as $\mathcal{R}_3$ can be obtained using well-known formulation techniques. Since it is desirable to announce whether a team has mathematically qualified for or eliminated from the playoff as soon as possible after a game is played (for example, the news media routinely announce this using the simple rule in a post-game show), the mixed integer program should be solved quickly, say in a few minutes. A practical criterion is that the mixed integer program should run using a solver's default settings as the task will be ran by a non-expert. Another interesting question is whether an open source free solver is competitive with a commercial solver in this setting.

Determining when a team is mathematically guaranteed to become a division leader for baseball can be solved by a network flow model proposed by Hoffman and Rivlin (1970). Since the hockey qualification is similar to baseball with a different number of wild cards[2], the same could be shown. In McCormick (1996) a fast flow based algorithm is given to determine when a team is eliminated from first place. The matter of additional teams being qualified based on their score, generally referred to as wild cards is more difficult. In Gusfield and Martel (2002) it is shown that when there are multiple divisions and wild cards (in our case three divisions and 5 wild cards in each conference) then determining whether a team is eliminated from the playoffs is *NP*-Hard.

Before giving a formulation for this problem we will describe some of the methods used by the press. The most common rule used by the press, or by curious fans, is simple and can be computed by hand. The rule goes as follows: to determine if a given Team A is qualified for the playoffs assume they lose all their remaining games and that every other team wins all their remaining games. Assign the points accordingly and then, based on the points, determine if Team A is qualified. If under this worst case scenario Team A qualifies for the playoffs, then clearly they have clinched a playoff spot. The problem with such a rule is that it is weak, as it may be impossible for all the other teams to win all their games, since only one team wins a given game. An example of a partially completed schedule where a team is qualified, yet whose qualification is unrevealed by this rule is easy to construct. When many teams are in a tight race for the last couple of playoff spots in a conference, this assumption goes too far. This rule is also slightly improved by tallying the wins of each team in addition to their score and using this to break ties when they occur.

Such a simple rule for determining teams does not give a false positive result, saying a team is qualified when they are not, but as mentioned it can fail to identify a qualified team, and as we shall show, this does occur in practice. In Section 2, we give a mixed integer program formulation that is theoretically better than the simple rule used by the media in determining when a team is mathematically qualified for or eliminated from the NHL playoffs. In Section 3, real data is used to see whether such formulation is of practical value, that is, does it really give better results than using the simple rule? Finally, we explore the difficulty of solving such MIPs using commercial and public domain solvers.

2. FORMULATION

2.1 Level 1 tie breaking

[2]The term wild card refers to a team that qualifies for the championship playoffs without winning its specific subdivision (usually called a conference or division) outright.

Here we give a formulation to detect qualification and elimination with respect to $\mathcal{R}_1$. This formulation will be used for each team on each day of the schedule. There will be one Mixed Integer Program that determines when a team is qualified for the playoffs and one that determines when they have been eliminated from the playoffs, a different MIP will be solved each day for each team to answer each of these questions. The feasible region of the polyhedron we define will consist of outcomes when a given team will be eliminated (qualified) from the playoffs. If we can show that either of these regions is empty, then we know that the team must be qualified (eliminated) from the playoffs.

Recall that $\mathcal{R}_1$ considers the scores and wins of all teams when breaking ties. Before giving the technical details of our formulation we state its goal: For any team k, to determine if it is mathematically qualified for the playoffs the MIP will be formulated so that any feasible solution gives an outcome for the future games that could possibly eliminate team k; thus, if no such outcome exists, team k is mathematically qualified. This can be updated on daily basis for each team and solved to determine which of all the teams are mathematically qualified. A similar formulation can be constructed to determine mathematical elimination.

This MIP is designed to answer the question "Can team k be eliminated?" If the answer to this question is no, then one can conclude that they are mathematically qualified, if the answer is yes, a feasible solution gives a completed schedule that could possibly[3] eliminate them. In answering this question, the formulation represents the outcome of future games, some assumptions are made to represent the MIP in a simplified but equivalent form. Since only feasibility is concerned, we can make some assumptions to shrink the dimension and complexity of our polyhedron without changing its feasibility. Since the question is "Can team k be eliminated?", the strategy is to fix all outcomes as bad as possible for team k. We therefore assume the following:

1. Team k loses all of its remaining games.
2. All teams in the same conference as team k win all of their remaining games with teams in the other conference.
3. All remaining games played between teams other than k will end in overtime losses.

The justification for these assumptions is straightforward, assuming k loses all of their remaining games can clearly be done, because we are searching for outcomes in which they can be eliminated. Assuming all teams in the same conference as k win their cross conference games can be done because this will only improve all the teams scores in the same conference as k without putting any of them at a disadvantage. Finally, assuming that games end with overtime losses will only boost all the scores of teams besides k, since the points in these games awarded would always be two points and one point, instead of two points and zero points.

[3]Note that such a configuration may not necessarily eliminate them

Qualification formulation for team *k*

maximize: 0 (1)

subject to: $t_k = s_k + \epsilon \times w_k$ (2)

$$t_i = s_i + 2 \times g_{ik} + \sum_{j \in C_k} \left(g_{ij} + x_{ij} \right) \quad (3)$$

$$+2 \times \sum_{j \notin C_k} (g_{ij}) + \epsilon \times \Big(w_i + g_{ik} \quad (4)$$

$$+ \sum_{j \in C_k} \left(g_{ij} + x_{ij} \right) + \sum_{j \notin C_k} (g_{ij}) \Big) \quad \text{for} \quad \text{for } i \neq k \quad (5)$$

$$x_{ij} + x_{ji} = g_{ij} \text{ for } i, j \in C_k \text{ and } i, j \neq k \quad (6)$$

$$t_i - t_j \leq M(1 - z_j) \text{ for } i, j \in C_k \text{ and in the same division} \quad (7)$$

$$\sum_{i \in D} z_i = 1 \quad \text{for each division } D \text{ in } C_k \quad (8)$$

$$\left(t_k + N \times z_k \right) - \left(t_i + N \times z_i \right) \leq M \times (1 - y_i) \text{ for } i \in C_k \quad (9)$$

$$\sum_{i \in C_k, i \neq k} y_i \geq 8 \quad (10)$$

integers: $x_{ij} \geq 0$ (11)

binary: z_i, y_i (12)

In this formulation we have ϵ small and $M >> N$ both large. The variable x_{ij} represents the number of times team i beats team j in their remaining unplayed games, $g_{ij} = g_{ji}$. In the formulation C_k represents the conference containing team k. The variables t_i for each i in the conference of team k represent the score of the team plus ϵ times the number of wins the team has in a completed schedule. This is computed using the numbers s_i and w_i, which represent the current score and wins of team i in the completed part of the schedule, along with points achieved from a possible completion of the remaining games in the schedule. This adjusted score t_i is then used to determine the value of the variables z_i for each team i in the conference. The binary variable $z_i = 1$ if and only if team i is the leader of its respective division. The binary variable $y_i = 1$ if and only if team i is ahead of or tied with team k in the standings. The definition of these variables and constants are summarized in Table 3. If at least eight such teams exist then we have a feasible schedule where team k can be eliminated. Note that if team k ties other teams we count team k as losing to avoid the possibility of falsely claiming team k is qualified. We again note the example where all teams tie all games and remark that further rules would need to be developed to break ties.

Given a partially completed schedule we can create this MIP for each team and determine if it is mathematically qualified. In order to determine when a team is mathematically eliminated from the playoffs a similar MIP can be used. We will not describe the elimination MIP here because it is very similar and answers the question "Can team k possibly qualify for the playoffs?" and if the answer to this is no, then they are eliminated.

Table 3. Variables and constants

s_i	Current score of team i	Constant
w_i	Current wins of team i	Constant
g_{ij}	Games remaining between i and j	Constant
t_i	Adjusted score of i given completed schedule	Variable
z_i	Indicates division leaders	Variable
y_i	Indicates team i's lead above team k	Variable
C_i	Conference containing team i	Variable
x_{ij}	Represents future wins of i against j	Variable

In "http://www.ilog.com/products/cplex/" the authors approached the qualification problem for Brazilian Football in a similar way but with one difference, instead of asking the question "Can team k be eliminated?" they sought to answer a more specific question "What is the largest number of points team k can earn and still be eliminated?" This is a stronger question, and if they have already achieved too many points to be eliminated, then it is concluded that they have qualified. The answer to such a feasible MIP gives a threshold value that a team can earn and be guaranteed a qualification slot, as opposed to our method which just gives the "no news yet" answer. We initially looked at answering this question but found it too computationally demanding to be of any use. In addition to adding an objective function, including this information requires us to scale back the assumptions we made in order to shrink the feasible region of our problem, giving us a much more computationally difficult problem to solve. We note that extending the formulation back to finding this threshold for each team is not difficult.

2.2 Level 3 tie breaking

In this section we present an extension of the first formulation that determines qualification with respect to $\mathcal{R}_3$. In addition to the previously included rules, the following is included: the differentiation between teams who have the same number of points and the same number of wins, by comparing the points, and then wins earned against this tied subset of the teams. Unfortunately, this more complicated and numerically demanding formulation is not of practical use by even the cutting edge commercial MIP solvers, nevertheless, it is still presented for theoretical interest. The modification of the existing formulation is to create adjusted scores b_{ti} for each team that includes an extra bonus to help differentiate between tied teams. In the following formulation the variable b_{ij} represents the small bonus score added to team i for points scored against team j in the event that both teams i and j are tied with team k. The variable v_{ij} is the value of the bonus that would be added to team i against team j, if j is tied with k, and thus b_{ij} is set to be equal to v_{ij} or 0. We also make sure that $\epsilon >> \epsilon_1 >> \epsilon_2$. The new variables s_{ij} and w_{ij} represent the number of points and the number of wins that team i has earned in games against team j, this additional information is needed in this formulation. The definition of these additional variables and constraints are summarized in Table 4. The constraints 23-26 are: $t_i \neq t_k \Rightarrow b_{ij} = 0\ \forall i, j \in C_k$, and the constraints 27-32 are $t_j = t_k \Rightarrow b_{ij} = v_{ij}$. These constraints are constructed using binary variables and standard integer programming techniques.

Table 4. Additional variables and constants

s_{ij}	Current score of team i against team j	Constant
w_{ij}	Current wins of team i against team j	Constant
t_i	Adjusted score of i with extra bonus	Variable
b_{ij} , v_{ij}	Used to compute extra bonuses	Variable
b_{ij}	Binary variables used to form implications	Variable

Qualification formulation for team k

$$\text{maximize: } 0 \tag{13}$$

$$\text{subject to: } t_k = s_k + \epsilon \times w_k \tag{14}$$

$$t_i = s_i + 2\times g_{ik} + \sum_{j\in C_k}\left(g_{ij} + x_{ij}\right) \tag{15}$$

$$+2\times \sum_{j\notin C_k}(g_{ij}) + \epsilon\times\left(w_i + g_{ik}\right. \tag{16}$$

$$\left. + \sum_{j\in C_k}\left(g_{ij}+x_{ij}\right) + \sum_{j\notin C_k}(g_{ij})\right) \text{ for } i \neq k \text{ for} \tag{17}$$

$$x_{ij} + x_{ji} = g_{ij} \text{ for } i, j \in C_k \text{ and } i, j \neq k \tag{18}$$

$$\hat{t}_i = t_i + \sum_{j\in C_k} b_{ij} \text{ for } i \in C_k \tag{19}$$

$$v_{ij} = \epsilon_1 \times\left(s_{ij} + g_{ij} + x_{ij}\right) + \epsilon_2\times\left(x_{ij} + w_{ij}\right) \tag{20}$$

$$\text{for } i, j \in C_k, i \neq j, \text{ and } j \neq k \tag{21}$$

$$v_{kj} = \epsilon_1 \times\left(s_{ij} + g_{ij} + x_{ij}\right) + \epsilon_2\times(w_{kj}) \text{ for } j \neq k \tag{22}$$

$$t_j - t_k \leq M\times(1-w_j) \text{ for } j \in C_k \tag{23}$$

$$t_k - t_j \leq M\times(1-w_j) \text{ for } j \in C_k \tag{24}$$

$$-b_{ij} \leq M\times w_p \text{ for } i, j \in C_k \tag{25}$$

$$b_{ij} \leq M\times w_p \text{ for } i, j \in C_k \tag{26}$$

$$(b_{ij} - v_{ij}) - (t_j - t_k) \leq M\times h_j^1 \text{ for } i, j \in C_k \tag{27}$$

$$(b_{ij} - v_{ij}) - (t_k - t_j) \leq M\times h_j^2 \text{ for } i, j \in C_k \tag{28}$$

$$(t_j - t_k) - (b_{ij} - v_{ij}) \leq M\times h_j^3 \text{ for } i, j \in C_k \tag{29}$$

$$(t_k - t_j) - (b_{ij} - v_{ij}) \leq M\times h_j^4 \text{ for } i, j \in C_k \tag{30}$$

$$h_j^1 + h_j^2 \leq 1 \text{ for } j \in C_k \tag{31}$$

$$h_j^3 + h_j^4 \leq 1 \text{ for } j \in C_k \tag{32}$$

$$\hat{t}_i - \hat{t}_j \leq M\times(1-z_j) \text{ for } i, j \in C_k \text{ in same division} \tag{33}$$

$$\sum_{i\in D} z_i = 1 \text{ for each division } D \text{ in } C_k \tag{34}$$

$$\left(\hat{t}_k + N \times z_k\right) - \left(\hat{t}_i + N \times z_i\right) \le M \times (1 - y_i) \text{ for } i \in C_k \quad (35)$$

$$\sum_{i \in C_k, i \ne k} y_i \ge 8 \quad (36)$$

integers: $x_{ij} \ge 0$ (37)

binary: z_i, y_i, w_i, b_j^i (38)

As in the previous case, the formulation for elimination is similar enough that we will not state it, it uses essentially the same ideas and methods.

3. COMPUTATIONAL RESULTS

These MIPs are used to determine qualification/ elimination with respect to R_1 on the data from the 2003-2004 NHL season, the most recent completed season (the 2004-2005 was canceled due to a labor dispute), and found it to be an improvement over rules used by the media described in the introduction. For most of the teams the simple rule was able to detect qualification and elimination at the same time as our method, but our method determined elimination of one team early and qualification of one team early. We were able to announce the qualification of the Colorado Avalanche on March 20th instead of March 22nd 2004 as was predicted by the simple rule, and were able to announce the elimination of the Anaheim Mighty Ducks on March 25th instead of March 26th 2004. All of the MIPs from this season were solved in a few minutes using CPLEX 9.1 and a Pentium 4 class machine. To see whether open source or other free solvers can solve these problems within a few minutes, we first tried solvers such as glpk 4.0 and lpsolve. However, they were too slow. For example, glpk took several hours to solve some instances and many instances were too big for lpsolve. This is perhaps an unfair comparison as these are mainly LP solvers without advanced MIP techniques built into the solvers. Next we tried free MIP solvers with an industrial strength MIP implementation. They are:

1. SCIP with Clp from COIN-OR as its LP solver,
2. Cbc with Clp as its LP solver from COIN-OR, and
3. SYMPHONY with Clp as its LP solver from COIN-OR.

In this setting, SCIP and Cbc solved all problems within a few minutes but SYMPHONY had problems solving some of the instances.

When we implemented the formulation for R_3 we ran into numerical difficulties and were unable to get reliable results. Even choosing ϵ_* as large as possible without invalidating the formulation, the precision required to solve such integer programs was too much for CPLEX to handle even on the highest settings available because of numerical difficulties. SCIP and Cbc also reported problem in this setting such as taking hours to solve one problem.

When a single team is doing very well (or poorly) the simple rule can detect qualification (elimination) accurately as early as possible. Our method outperforms the simple rule when there were several teams with close scores fighting for the last couple of playoff spots. This is due to the fact that, unlike our mathematical model, the simple rule is unable to consider the large number of outcomes for the remaining games and their impact on playoff qualification (elimination). The simple rule assumes large group of teams win (lose) all their remaining games and this does not give the strongest result. Our source code and season data are available upon request.

4. CONCLUSION

It is now time to consider the answer to the questions stated at the beginning of this paper. We were able to give a mixed integer programming formulation for the mathematical qualification and elimination problem for the NHL playoffs that captures the first level of tie breaking rules. Our formulation was able to outperform the more conventional methods for determining qualification and elimination used by the press today. Finally, the instances of our problems were within the reach of the high powered commercial software CPLEX 9.1 as well as industrial strength free MIP solver like SCIP and Cbc. However, some other less powerful alternatives, such as glpk and lpsolve, proved too slow for practical use. The second formulation considered deeper tie breaking rules, but it was too complex to be solved because it required a higher level of precision than is currently available. We conclude that the methods presented here would be computationally impractical in the not so distant past.

Our work shows that there is room for improvement in the way the press calculates qualification and elimination of teams in major league sports. Despite the theoretical difficulty of this and similar problems, using current methods these problems can now be readily solved. Similar results have been achieved for other sports (Adler et al. (2002), Hoffman and Rivlin (1970), Ribeiro and Urrutia (2005)). Using these more sophisticated tools and software, new accuracy can be achieved in determining mathematical qualification and elimination in many sports, and due to the scale and interest in professional sports media outlets could provide the public with more accurate and timely information to that regard.

ACKNOWLEDGMENT

The authors would like to thank Professor Serge Kruk for help in running SCIP and Cbc. The second author would like to thank the Oakland University Honors College for its support through the Honors College Colloquium Award which supported this research while he was an undergraduate student.

REFERENCES

1. Adler, I., Erera, A., Hochbaum, D., and Olinich, E. (2002). Baseball, optimization and the world wide web. *Interfaces*, 32: 12-22.
2. Gusfield, D. and Martel, C. (2002). The structure and

complexity of sports elimination numbers. *Algorithmica*, 32: 73-86.
3. Hoffman, A. and Rivlin, T. (1970). When is a team "mathematically" eliminated? In: H.W. Kuhn (Ed.), *Princeton Symposium on Mathematical Programming*, Princeton University Press, Princeton, NJ, pp. 391-401.
4. McCormick, T. (1996). Fast algorithms for parametric scheduling come from extensions to parametric maximum flow. *Proceedings of the 28th Annual ACM Symposium on Theory of Computing*, Philadelphia, United States, pp. 394-422.
5. Ribeiro, C.C. and Urrutia, S. (2005). An application of integer programming to playoff elimination in football championships. *International Transactions in Operational Research*, 12: 375–386.
6. http://scip.zib.de/.
7. https://projects.coin or.org/Cbc.
8. https://projects.coin or.org/Clp.
9. https://projects.coin or.org/SYMPHONY.
10. http://www.coin or.org/index.html.
11. http://www.cs.sunysb.edu/algorith/implement/lpsolve/implement.shtml.
12. http://www.gnu.org/software/glpk/.
13. http://www.ilog.com/products/cplex/.