6
Os Problemas de Classificação: FutMax

O Campeonato Brasileiro é o torneio de futebol mais importante no Brasil, e possivelmente, o maior no mundo em termos do número de equipes participamtes. É acompanhado por milhões de pessoas, que assistem aos jogos nos estádios, acompanham as transmissões de rádio e televisão e procuram nos jornais, no rádio, na televisão e, mais recentemente, na Internet informações acerca do desempenho e das chances de suas equipes favoritas.

As características do torneio que se expõem a seguir correspondem às do Campeonato Brasileiro de 2002. O torneio está organizado em duas etapas: a etapa de classificação e a etapa eliminatória final ou simplesmente finais. Durante a etapa de classificação, cada equipe joga contra todas as demais equipes uma única vez. As equipes obtêm três pontos por cada vitória e um ponto por cada empate. As equipes são ordenadas em uma tabela de classificação pela quantidade de pontos obtida até o momento. Ao final da etapa de classificação, as oito primeiras equipes nessa tabela se classificam para a fase eliminatória final. As quatro últimas equipes nessa mesma tabela são rebaixadas de categoria e não jogam no campeonato do ano seguinte. Em 2002 este campeonato foi disputado entre agosto e dezembro por 26 das principais equipes do Brasil. Os 325 jogos da etapa de classificação foram divididos em 29 rodadas de 13 ou menos jogos cada uma. As rodadas não eram completas e algumas equipes folgavam em algumas rodadas.

O futebol é uma importante atividade econômica no Brasil. Uma equipe que não se classifica para as finais perde muito dinheiro e pode ser ainda forçada a vender seus melhores jogadores para cobrir suas despesas de manutenção durante este período, quando perde toda a arrecadação por venda de ingressos e direitos televisivos. A imprensa e os torcedores especulam e acompanham a tabela de classificação todos os dias.

O objetivo principal de todas as equipes é de se classificar entre as oito primeiras equipes na tabela de classificação ao final da primeira fase. Para as equipes que não podem alcançar este objetivo, seu segundo objetivo é, pelo menos, não terminar nas últimas quatro posições para se manterem na competição no ano seguinte.

A imprensa oferece várias estatísticas para ajudar aos torcedores a avaliar o desempenho das suas equipes favoritas. Uma das informações é chamada de "chance de classificação". Este número fornece, para cada equipe, a probabilidade de se classificar entre os oito primeiros ao final da primeira etapa. Embora sem apresentar um modelo probabilístico bem definido, o cálculo destas probabilidades está baseado no aproveitamento de cada equipe, isto é, na fração de pontos ganhos em relação ao número de pontos disputados. Desde as primeiras rodadas da primeira etapa, a imprensa informa que qualquer equipe estará classificada se conseguir atingir um determinado número de pontos. A estimação de este valor é uma aproximação baseada nas estatísticas correntes e em informações históricas.

Quando uma equipe atinge este número de pontos, a imprensa anuncia que a suas chances de classificação são de 100% e que a equipe está "matematicamente classificada". Freqüentemente, esta informação é incorreta. Por exemplo, este valor foi avaliado em 41 pontos em 2002. A equipe do São Paulo atingiu 43 pontos na rodada 24 no dia 31 de outubro e a imprensa anunciou que o São Paulo estava "matematicamente classificado". Entretanto, o modelo desenvolvido neste tese e apresentado a seguir foi capaz de provar facilmente que esta informação não era verdadeira, mostrando uma combinação de resultados dos jogos restantes que deixavam o São Paulo fora das oito primeiras posições. Por mais que a probabilidade dessa combinação de resultados fosse pequena, seria uma surpresa e uma situação difícil para a imprensa se o São Paulo não fosse classificado ao final da primeira etapa do campeonato.

O primeiro problema que se aborda neste capítulo da tese é o *Problema da Classificação Garantida* (PCG). Este problema consiste em calcular o *Número de Pontos para Classificação Garantida* (NCG), que é o menor número inteiro tal que se uma determinada equipe obtém NCG ou mais pontos ela garante a sua classificação, independentemente do desempenho das outras equipes. Este número depende do número de pontos de cada equipe no momento do cálculo e dos jogos que ainda serão disputados. O segundo problema abordado é o *Problema da Classificação Possível* (PCP), que consiste em calcular o *Número de Pontos para Classificação Possível* (NCP). Este é o menor número de pontos que uma determinada equipe precisa obter para ter alguma chance de se classificar. Este número também depende do número de pontos de cada equipe no momento do cálculo e dos jogos que ainda serão disputados.

Nota-se que $NCP \leq NCG$ para cada equipe, a qualquer momento da etapa de classificação. O valor de NCG para uma equipe qualquer nunca pode aumentar de uma rodada para outra, enquanto o valor de NCP nunca pode diminuir. Obter um mínimo de NCG (resp. NCP) pontos é uma condição suficiente (resp. necessária) para a classificação. Uma equipe está *matematicamente classificada* para as finais se e somente se o número de pontos que já obteve, é maior ou igual ao seu NCG. Somente neste ponto a sua classificação pode ser anunciada sem nenhuma possibilidade de erro. Analogamente, uma equipe pode ser dita *matematicamente desclassificada* quando o número total de pontos que ainda tem que disputar (i.e., o número de jogos restantes multiplicado por três) mais o número de pontos já obtidos (MNP) é menor que seu NCP. As formulações usadas na solução destes problemas podem ser facilmente adaptadas para lidar com o segundo objetivo, qual seja, o de não terminar nas quatro últimas posições na tabela de classificação. Simplesmente, se consideram os números de pontos, necessários e suficientes, para ficar entre as 22 primeiras equipes.

Este capítulo está organizado da seguinte maneira. Na próxima seção, estuda-se a complexidade dos problemas de classificação e é revisada a bibliografia sobre problemas similares em esportes como basebol e basquete. A formulação geral dos problemas de classificação é discutida na Seção 6.2. Os modelos de programação inteira associados aos problemas PCG e PCP são introduzidos na Seção 6.3. Resultados numéricos, assim como diversos aspectos da aplicação dos modelos ao Campeonato Brasileiro 2002, são discutidos na Seção 6.4. As conclusões encerram o capítulo na última seção.

6.1 Complexidade

Para ilustrar o Problema da Classificação Garantida, considera-se um pequeno exemplo no qual três equipes competem pelas duas últimas vagas nas finais. As três equipes são Flamengo, Cruzeiro e Bahia, que têm respectivamente 37, 37 e 36 pontos na tabela de classificação:

Equipe	Pontos
Flamengo	37
Cruzeiro	37
Bahia	36

O Flamengo ainda tem que jogar contra Vasco e Bahia nas últimas duas rodadas, o Cruzeiro contra Grêmio e Santos e o Bahia contra Flamengo e Fluminense. A pergunta é: "qual é o menor número de pontos tal que se o Cruzeiro obtém este número de pontos ou mais ele garante a sua classificação em uma das vagas restantes?" É fácil ver que no caso de obter mais seis pontos (ganhando seus dois últimos jogos) o Cruzeiro garante a sua classificação. Se o Cruzeiro obtiver mais quatro pontos, ele terá 41 pontos ao final da etapa de classificação. Dependendo do resultado do jogo entre Flamengo e Bahia, existem três possíveis cenários para as possíveis melhores pontuações dessas duas equipes: Flamengo 43 e Bahia 39 (o Flamengo vence o Bahia, que é eliminado), Flamengo 41 e Bahia 40 (o jogo entre Flamengo e Bahia acaba empatado e o último é eliminado) ou Flamengo 40 e Bahia 42 (o Bahia vence o Flamengo, que é eliminado). O Cruzeiro estará classificado para as finais em qualquer destes cenários. Considera-se agora o que acontece se o Cruzeiro ganha um jogo e perde o outro, obtendo 40 pontos. Neste caso, se o Flamengo vence o Vasco e o Bahia ganha seus dois jogos restantes, o Bahia terá 42 pontos, o Flamengo 40 e o Cruzeiro também 40. Flamengo e Cruzeiro acabariam empatados. O Cruzeiro poderia não se classificar, dependendo da regra de desempate aplicada. Portanto, $NCG = 41$ para o Cruzeiro neste caso. Este pequeno exemplo ilustra o tipo de situação que pode acontecer para um número pequeno de equipes disputando a classificação. O problema é certamente mais complexo e difícil de analisar quando há muitas equipes competindo por vagas nas finais.

Durán [19] mostrou que o NCG não é equivalente ao mínimo numero de pontos que uma equipe precisa fazer para garantir a sua classificação. Existem situações nas quais uma equipe garante a sua classificação obtendo p pontos, mas não a garante obtendo $p+1$ pontos. Considere-se, por exemplo, uma situação na qual as mesmas três equipes competem pela última vaga nas finais com a seguinte tabela de classificação:

Equipe	Pontos
Cruzeiro	40
Flamengo	37
Bahia	37

O Flamengo ainda tem que jogar contra Cruzeiro e Vasco nas últimas duas rodadas, o Cruzeiro contra Flamengo e Bahia e o Bahia contra Fluminense e Cruzeiro. Observa-se facilmente que o Cruzeiro estará classificado obtendo quatro ou mais pontos nas duas últimas rodadas (totalizando pelo menos 44 pontos), já que nem Flamengo nem Bahia podem obter mais do que 43 pontos. O Cruzeiro não garante a sua classificação obtendo três pontos nas últimas duas rodadas, porque neste caso perderia o jogo contra o Flamengo ou o jogo contra o Bahia. A equipe que o vencer poderia ganhar também o jogo restante, ficando empatada com o Cruzeiro em 43 pontos. Portanto, $NCG = 44$ para o Cruzeiro neste caso.

Considere-se agora o que acontece se o Cruzeiro obtém mais dois pontos, empatando os jogos contra o Flamengo e contra o Bahia e chegando aos 42 pontos. A máxima pontuação possível para Flamengo e Bahia neste caso é 41 pontos, dado que ao empatar com o Cruzeiro a máxima quantidade de pontos que podem obter nas últimas duas rodadas é quatro. Portanto, neste caso, a mínima quantidade de pontos com a qual o Cruzeiro garante a sua classificação é 42 e este número é diferente do seu NCG, que é 44.

Quase todo o trabalho prévio que trata deste tipo de problemas concerne à *Major League Baseball* (MLB) nos Estados Unidos. Segundo a notação de Gusfield e Martel [32], diz-se que a MLB segue a regra $\{(1,0)\}$: dado que um jogo não pode terminar em empate, uma das equipes necessariamente obterá um ponto e a outra zero. O Campeonato Brasileiro segue a regra $\{(3,0),(1,1)\}$: se uma equipe vence ele ganha três pontos e a outra zero, se o jogo acabar empatado ambas equipes ganham um ponto. Os problemas de classificação são mais fáceis para torneios que seguem a regra $\{(1,0)\}$ do que para aqueles que seguem a regra $\{(3,0),(1,1)\}$ [6]. Uma outra característica que faz os cálculos para a MLB mais simples do que para o Campeonato Brasileiro, é que até alguns poucos anos atrás, o único objetivo na MLB era terminar na primeira posição de pequenos grupos de no máximo seis equipes, o que leva a muitas menos combinações de resultados possíveis.

Schwartz [56] resolveu o problema de determinar se uma equipe está eliminado na MLB com um algoritmo de fluxo máximo. Também para a MLB, Hoffman e Rivlin [35] descreveram condições necessárias e suficientes para que uma equipe esteja eliminada das k primeiras posições. Robinson [54] apresentou um modelo de programação linear inteira para calcular a diferença máxima de pontos que uma equipe pode atingir com respeito à segunda colocada. McCormick [40] mostrou que decidir se uma equipe está eliminada das primeiras k posições é NP-completo. Adler et al. [1] mostraram como a programação inteira pode ser usada para calcular o *Número de Pontos para Classificação Garantida* e o *Número de Pontos para Classificação Possível* para a MLB. Wayne [64] mostrou que em todo momento na MLB, existe um número de pontos tal que toda equipe está eliminada da primeira posição se e somente se não puder atingir esse número de pontos ao final do campeonato. Gusfield e Martel [32] generalizaram o resultado de Wayne, mostrando que esse número de pontos existe para todo campeonato com certas características, incluindo os de futebol que seguem a regra $\{(3,0),(1,1)\}$. Bernholt et al. [6] mostraram que o problema de decidir se uma equipe está eliminada do primeiro lugar para campeonatos de futebol com a regra $\{(3,0),(1,1)\}$ é NP-completo. Por conseguinte, determinar a eliminação das primeiras k posições é NP-completo (e em conseqüência, PCP também).

6.2 Formulação dos problemas

Dados n, a quantidade de equipes no campeonato ($n = 26$ para o Campeonato Brasileiro de 2002, aqui representado por CB02); $t \in 1,2,\ldots,n$, uma equipe específica; R, um conjunto de duplas, indicando as possíveis repartições de pontos como conseqüência do resultado de um jogo (por exemplo, $R = \{(3,0),(1,1),(0,3)\}$ para CB02, isto é, se alguma das equipes vence, ela recebe 3 pontos e o perdedor 0, e se o jogo termina empatado ambas as equipes obtêm 1 ponto); um vetor s de n elementos, onde s_i indica a quantidade de pontos já acumulados pela equipe i a qualquer momento do campeonato; e g, uma matriz quadrada e simétrica de dimensão n, onde g_{ij} indica a quantidade de jogos entre as equipes i e j que ainda devem ser jogados (no contexto do CB02, $g_{ij} = 0$ caso as equipes i e j já tenham se enfrentado, ou $g_{ij} = 1$ caso contrário). Definem-se:

Problema da Classificação Possível (PCP): Determinar o número mínimo de pontos que a equipe t deve fazer para poder classificar-se entre as k primeiras.

Problema da Classificação Garantida (PCG): Determinar o número mínimo de pontos tal que se a equipe t obtém essa quantidade de pontos ou mais, ela garante a sua classificação entre as k primeiras.

Uma *atribuição válida* é um conjunto de $n \cdot (n-1)$ triplas $A(i,j) = (p_1(i,j), p_2(i,j), p_3(i,j))$ de inteiros não negativos associadas a cada um dos pares ordenados (i,j) de equipes diferentes, tal que $p_1(i,j) + p_2(i,j) + p_3(i,j) = g_{ij}$, $p_1(i,j) = p_3(j,i)$ e $p_2(i,j) = p_2(j,i)$, onde $p_1(i,j)$, $p_2(i,j)$ e $p_3(i,j)$ representam, respectivamente, o número possível de vitórias da equipe i sobre a equipe j, o número possível de empates entre as equipes i e j e o número possível de vitórias da equipe j sobre a equipe i entre os g_{ij} jogos restantes entre elas.

Dada uma atribuição válida, o número total de pontos obtido por cada equipe $i = 1, \ldots, n$ ao final do campeonato é:

$$p_i = s_i + \sum_{j \neq i} 3 \cdot p_1(i,j) + \sum_{j \neq i} p_2(i,j).$$

Para toda atribuição válida no contexto do Problema da Classificação Garantida, a posição final da equipe $i = 1, \ldots, n$ na tabela de classificação é definida como $P_i = |\{j : 1 \leq j \leq n, j \neq i, p_j \geq p_i\}| + 1$. Então, o PCG consiste em encontrar, para cada equipe $t \in 1, 2, \ldots, n$, o menor inteiro NCG^t tal que para toda atribuição válida, se $p_t \geq NCG^t$ então $P_t \leq k$.

Similarmente, no contexto do Problema da Classificação Possível, a posição final da equipe $i = 1, \ldots, n$ na tabela de classificação é definida como $P_i = |\{j : 1 \leq j \leq n, j \neq i, p_j > p_i\}| + 1$. Então, o PCP consiste em encontrar, para cada equipe $t \in 1, 2, \ldots, n$, o menor inteiro NCP^t tal que existe uma atribuição válida onde $p_t = NCP^t$ e $P_t \leq k$.

A definição da posição de uma equipe na tabela de classificação é diferente no contexto dos problemas PCG e PCP. No caso do PCG, para assegurar que a equipe t está classificada qualquer que seja o critério de desempate, qualquer outra equipe com a mesma pontuação que t é considerada como melhor classificada do que esta. Entretanto, no contexto do PCP, somente é considerada a possibilidade (não a certeza) de classificação. Neste caso, a equipe t tem chances de se classificar ainda que esteja empatada com outras equipes na última vaga. Em conseqüência, qualquer outra equipe com a mesma pontuação que t é considerada como pior classificada do que esta.

6.3 Modelos de programação inteira

A notação definida na seção anterior é usada para modelar o Problema de Classificação Garantida como um problema de programação linear inteira.

Para cada $t = 1, \ldots, n$, seja $\underline{NCG^t}$ o máximo número de pontos tal que existe uma atribuição válida para a qual $p_t \geq \underline{NCG^t}$ e $P_t > k$. Então, $\underline{NCG^t}$ é o maior número de pontos que a equipe t pode obter e ainda ficar desclassificada. Portanto, $NCG^t = \underline{NCG^t} + 1$ é o número mínimo de pontos tal que se a equipe t obtém essa quantidade de pontos ou mais, ela garante a sua classificação entre as k primeiras. Definem-se as seguintes variáveis:

$$x_{ij} = \begin{cases} 1, & \text{se a equipe } i \text{ ganha da equipe } j \\ 0, & \text{caso contrário;} \end{cases}$$

$$y_j = \begin{cases} 1, & \text{se } p_j \geq p_t \text{ (i.e., se a equipe } t \text{ não está à frente da equipe } j) \\ 0, & \text{caso contrário.} \end{cases}$$

As variáveis x_{ij} são binárias porque no CB02 as equipes se enfrentam apenas uma vez. Em um modelo mais geral, as variáveis x_{ij} poderiam assumir qualquer valor inteiro positivo em um intervalo mais amplo. As variáveis x_{ij} e x_{ji} servem para modelar o resultado do jogo entre as equipes i e j ($x_{ij} = 1$ se a equipe i vence a equipe j, $x_{ji} = 1$ se a equipe j vence a equipe i e $x_{ij} = x_{ji} = 0$ se as equipes i e j empatam). Assume-se que $R = \{(3,0),(1,1),(0,3)\}$.

O seguinte modelo de programação linear inteira calcula $\underline{NCG}^t$, para $t = 1,\ldots,n$, i.e., o máximo número de pontos a equipe t pode obter e ainda ficar desclassificada:

$$\text{PCG}(t): \begin{cases} \underline{NCG}^t = & \text{máximo } p_t & \\ \text{sujeito a:} & & \\ x_{ij} + x_{ji} \leq g_{ij} & \forall\, 1 \leq i < j \leq n & (1) \\ p_j = s_j + 3\sum_{i \neq j} x_{ji} + \sum_{i \neq j}[g_{ij} - (x_{ij} + x_{ji})] & & \\ & \forall\, 1 \leq j \leq n & (2) \\ p_t - p_j \leq M(1 - y_j) & \forall\, 1 \leq j \leq n, j \neq t & (3) \\ \sum_{j \neq t} y_j \geq k & & (4) \\ x_{ij} \in \{0,1\} & \forall\, 1 \leq i \leq n, 1 \leq j \leq n, i \neq j & \\ y_j \in \{0,1\} & \forall\, 1 \leq j \leq n\,, j \neq t. & \\ p_j \geq 0 & \forall\, 1 \leq j \leq n. & \end{cases}$$

A função objetivo consiste em determinar o maior número de pontos $\underline{NCG}^t$ que a equipe t pode fazer e ainda ficar desclassificada. As restrições (1) determinam que só uma equipe pode ganhar cada jogo. As restrições (2) estabelecem o número total de pontos obtidos por cada equipe ao final do campeonato. Para cada equipe $j = 1,\ldots,n$, a soma $\sum_{i \neq j} x_{ji}$ dá o número de vitórias que a equipe j obtém nos jogos ainda não disputados (três pontos por cada vitória), enquanto $\sum_{i \neq j}[g_{ij} - (x_{ij} + x_{ji})]$ corresponde ao número de empates obtidos pela equipe j nos jogos ainda não disputados (um ponto por cada empate).

Seja M um limite superior para a maior diferença possível entre o número de pontos obtidos por qualquer par de equipes. Dado que no CB02 participaram 26 equipes e cada uma jogava uma única vez com cada uma das demais, $M \geq 3 \cdot 25 = 75$ é um limite superior para $p_t - p_j$ para qualquer par (t, j) de equipes com $j, t = 1, \ldots, n$ e $j \neq t$. As restrições (3) estabelecem que se $p_j < p_t$, então $y_j = 0$ (i.e., a equipe t está na frente da equipe j na tabela de classificação). A restrição (4) força a equipe t a não estar classificado entre as k primeiras ($k = 8$ para o CB02).

$PCG(t)$ determina a quantidade máxima de pontos que a equipe t pode fazer e ainda assim ficar eliminada das k primeiras posições. Então, pode se dizer que a equipe t está matematicamente classificada assim que $PCG(t)$ torna-se inviável. Quanto mais ajustado for o limite superior M, mais rapidamente será obtida a solução do modelo por qualquer resolvedor de problemas de programação linear inteira. Um limite superior de $p_t - p_j$ mais justo é dado por $M_j = s_t - s_j + 3\sum_{i \neq t} g_{it}$. Este limite é usado na implementação.

Considera-se agora a formulação como um problema de programação linear inteira do Problema da Classificação Possível para uma equipe $t = 1, \ldots, n$. Seja NCP^t o menor número de pontos tal que existe pelo menos uma atribuição válida para a qual $p_t = NCP^t$ e $P_t \leq k$. Definem-se as seguintes variáveis adicionais:

$$z_j = \begin{cases} 1, & \text{if } p_j > p_t \text{ (i.e., se a equipe } j \text{ está à frente da equipe } t) \\ 0, & \text{caso contrário.} \end{cases}$$

O modelo anterior pode ser reformulado para resolver este novo problema:

$$\text{PCP}(t): \begin{cases} NCP^t = & \text{mínimo } p_t & \\ \text{sujeito a:} & & \\ x_{ij} + x_{ji} \leq g_{ij} & \forall\, 1 \leq i < j \leq n & (1) \\ p_j = s_j + 3\sum_{i \neq j} x_{ji} + \sum_{i \neq j}[g_{ij} - (x_{ij} + x_{ji})] & & \\ & \forall\, 1 \leq j \leq n & (2) \\ p_j - p_t \leq M z_j & \forall\, 1 \leq j \leq n, j \neq t & (3') \\ \sum_{j \neq t} z_j \leq k - 1 & & (4') \\ x_{ij} \in \{0, 1\} & \forall\, 1 \leq i \leq n, 1 \leq j \leq n, i \neq j & \\ z_j \in \{0, 1\} & \forall\, 1 \leq j \leq n\,, j \neq t. & \\ p_j \geq 0 & \forall\, 1 \leq j \leq n & \end{cases}$$

As restrições (3') cumprem o mesmo papel que as restrições (3) na formulação $PCG(t)$. Obrigam que se $p_j > p_t$, então $z_j = 1$ (i.e., a equipe j está melhor classificada que a equipe t). Neste caso, o valor M_j que substitui M é dado por $M_j = s_j - s_t + 3\sum_{i\neq j} g_{ij}$. A restrição (4') estabelece que existem no máximo $k-1$ equipes melhor classificadas que a equipe t na tabela de classificação. A inviabilidade de $PCP(t)$ significa que a equipe t está matematicamente eliminada, i.e., não pode mais se classificar para as finais.

Estes modelos podem ser estendidos para considerar alguns dos mais usuais critérios de desempate. No caso do Campeonato Brasileiro 2002 o primeiro critério de desempate era o número de vitórias obtidas: para todo par de equipes com a mesma quantidade de pontos, considerava-se melhor classificada aquela que tiver obtido mais vitórias. Uma pequena modificação nos modelos pode tratar este critério: substitui-se as restrições (2) por $p_j = s_j + (3+\epsilon)\sum_{i\neq j} x_{ji} + \sum_{i\neq j}[g_{ij} - (x_{ij} + x_{ji})]$. Com está modificação, cada vitória dá $3+\epsilon$ pontos e não apenas três. Dado que cada equipe faz 25 jogos, se $\epsilon < 1/25$ a ordem determinada pelo número de pontos não é afetada por esta modificação. Entretanto, se duas equipes obtém a mesma quantidade de pontos, aquela com mais vitórias estará melhor classificada do que a outra.

6.4 Resultados

O Campeonato Brasileiro 2002 começou no 10 de agosto. Os problemas PCP e PCG foram considerados em dois contextos. No primeiro, considerou-se a classificação entre as oito primeiras posições, que definiam as oito equipes que classificavam para as finais. No segundo caso, considerou-se a não classificação entre as quatro últimas posições, que definiam as equipes que seriam rebaixadas de categoria no ano seguinte. Não se classificar para as finais significa uma grande perda, pois os jogos das finais são os de maior interesse do público e com direitos de transmissão mais caros. Classificar-se entre os quatro últimos equipes traz enormes dificuldades para estas equipes, que provavelmente deverão desmantelar suas estruturas devido ao rebaixamento e à menor importância dos jogos da segunda divisão.

Os modelos $PCG(t)$ e $PCP(t)$ foram resolvidos imediatamente após cada rodada, para cada equipe $t = 1, \ldots, n$. Dado que 26 equipes participavam do campeonato e se pretendia obter os valores de NCG e de NCP nos dois contextos descritos acima (classificação para as finais, com $k = 8$,

Rodada	Tempo (s)	Rodada	Tempo (s)	Rodada	Tempo (s)
1	871.2	11	14.6	21	0.9
2	856.8	12	16.0	22	0.7
3	743.4	13	15.6	23	0.7
4	697.0	14	15.3	24	0.7
5	710.3	15	1.3	25	0.6
6	677.8	16	4.7	26	0.5
7	757.5	17	2.9	27	0.1
8	457.1	18	2.6	28	<0.1
9	888.6	19	2.1	29	<0.1
10	39.7	20	2.1		

Tabela 6.1: Tempos de processamento para calcular NCG.

e não rebaixamento, com $k = 22$), 104 problemas de programação linear inteira eram resolvidos imediatamente após cada rodada. Usou-se o CPLEX 5.0 como o resolvedor de problemas de programação linear inteira. Os valores NCG e NCP para cada equipe, além de outras estatísticas, eram imediatamente disponibilizados no sítio web do projeto *FutMax* [48]. Os tempos de processamento são altos nas primeiras rodadas do campeonato. Entretanto, após algumas rodadas, e sobretudo perto do final da etapa de classificação, quando o interesse pelos resultados dos modelos é máximo, os tempos de processamento são de uns poucos segundos em um processador Pentium IV padrão. A Tabela 6.1 mostra os tempos de processamento (em segundos) para computar NCG para o Fluminense após cada rodada.

Diversas matérias sobre o projeto *FutMax* foram feitas pela imprensa brasileira (jornais, rádio e televisão) desde 2002. Nestas matérias, sempre era discutido o uso destes números pela própria imprensa.

Os modelos apresentados na Seção 6.3 e o sítio web do projeto *FutMax* mostraram sua utilidade várias vezes durante o campeonato. Após a décima primeira rodada, no dia 15 de setembro de 2002, o treinador do Vasco da Gama, Antonio Lopes, disse em conferência de imprensa que sua equipe estaria classificada para as finais se conseguisse ganhar dez dos 14 jogos restantes. O trabalho desenvolvido dentro do projeto *FutMax* permitiu mostrar que esta afirmação não era verdadeira e que existia uma combinação de resultados dos jogos restantes que deixavam o Vasco da Gama fora da zona de classificação ainda que ganhasse todos seus jogos restantes.

O São Paulo atingiu 43 pontos após a rodada 24 no dia 31 de outubro de 2002. Usando estimações de desempenho, a imprensa anunciou que esta equipe era a primeira matematicamente classificada para as finais. Novamente, a resolução dos modelos mostrou que isto não era verdade. A Tabela 6.2 (a) mostra a tabela de classificação após a rodada 24, enquanto que a Tabela 6.2 (b) ilustra um possível cenário para a tabela de classificação ao final da etapa de classificação, pela qual o São Paulo ficaria fora das oito primeiras posições e teria sido desclassificado.

	Equipe	Pt	Vt
1	**São Paulo**	43	13
2	São Caetano	39	12
3	Corinthians	38	11
4	Juventude	37	11
5	Santos	36	10
6	Atlético MG	33	10
7	Grêmio	33	9
8	Fluminense	31	9
9	Coritiba	30	9
10	Ponte Preta	30	9
11	Goiás	30	8
12	Vitória	28	8
13	Guarani	28	8
14	Figueirense	27	8
15	Atlético PR	27	7
16	Portuguesa	26	7
17	Bahia	25	7
18	Internacional	25	6
19	Cruzeiro	24	6
20	Vasco	23	7
21	Paraná	23	7
22	Palmeiras	23	5
23	Paysandú	22	7
24	Gama	22	6
25	Flamengo	22	6
26	Botafogo	22	5

(a)

	Equipe	Pt	Vt
1	Corinthians	48	14
2	Fluminense	46	14
3	Juventude	46	14
4	Coritiba	45	14
5	Atlético MG	45	14
6	São Caetano	45	14
7	Grêmio	45	13
8	Santos	45	13
9	**São Paulo**	45	13
10	Guarani	40	12
11	Vitória	38	11
12	Ponte Preta	32	9
13	Goiás	32	8
14	Figueirense	29	8
15	Vasco	29	8
16	Portuguesa	29	7
17	Atlético PR	29	7
18	Cruzeiro	28	6
19	Internacional	27	6
20	Paraná	26	7
21	Bahia	26	7
22	Paysandú	25	7
23	Flamengo	25	6
24	Palmeiras	25	5
25	Botafogo	25	5
26	Gama	23	6

(b)

Tabela 6.2: (a) A tabela de classificação após a rodada 24 e (b) uma possível tabela de classificação ao final da etapa de classificação.

As Figuras 6.1, 6.2 e 6.3 mostram gráficos ilustrando a evolução do NCG e NCP, no contexto da classificação para as *finais*, para três equipes. Além desses números, também é mostrada a evolução do número de pontos (P) acumulados após cada rodada e o máximo número de pontos (MNP) que a equipe poderia alcançar se ganhasse todos seus jogos restantes (três pontos adicionais por cada um). NCP e P nunca podem diminuir ao longo do campeonato, enquanto que NCG e MNP nunca podem aumentar. Uma equipe está matematicamente classificada para as finais assim que $P \geq NCG$. Analogamente, uma equipe está matematicamente eliminada das finais quando $MNP < NCP$. Uma equipe depende apenas de si mesma para se classificar para as finais se $MNP \geq NCG$.

A Figura 6.1 ilustra que MNP tornou-se menor do que NCG para o Fluminense na rodada 11, mostrando que nesse momento o Fluminense não dependia de si mesmo: ainda que obtivesse o maior número de pontos possível ganhando todos seus jogos restantes, existia pelo menos uma combinação de resultados que deixava o Fluminense fora da zona de classificação. Na rodada 25, o número de pontos acumulados pelo Fluminense atingiu seu NCP, significando que tinha uma chance de classificar-se ainda que perdesse todos os jogos restantes.

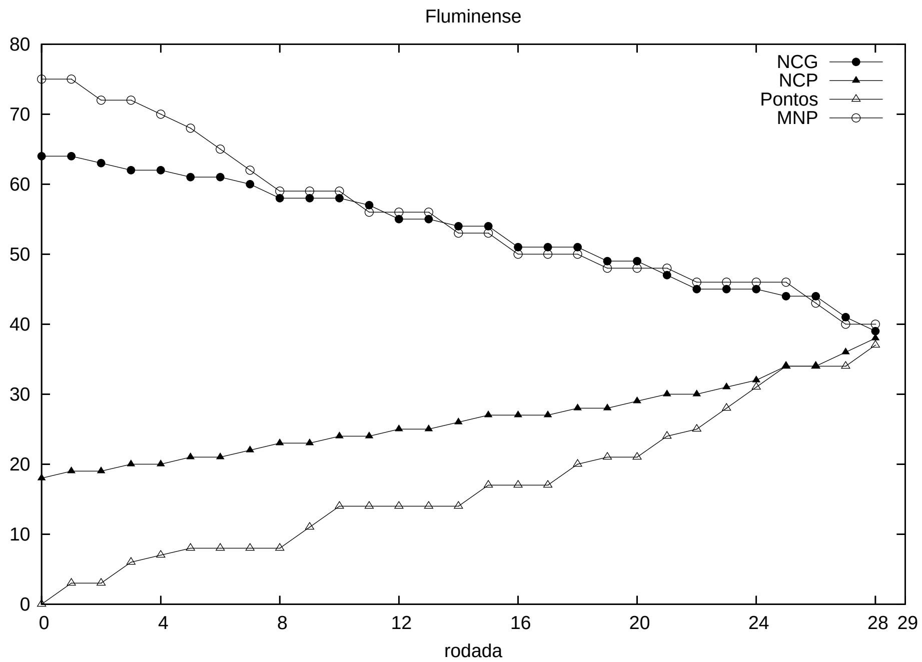

Figura 6.1: Fluminense no Campeonato Brasileiro 2002.

A Figura 6.2 ilustra os números do Palmeiras. Esta equipe foi eliminada das finais e foi uma das equipes classificados nas últimas quatro posições, não participando da edição 2003 do campeonato. O gráfico da Figura 6.3 mostra a evolução dos mesmos números para o São Caetano. Esta equipe sempre dependeu de si mesma e garantiu sua vaga nas finais após a rodada 26 (erradamente, a imprensa anunciou sua classificação matemática após a rodada 25).

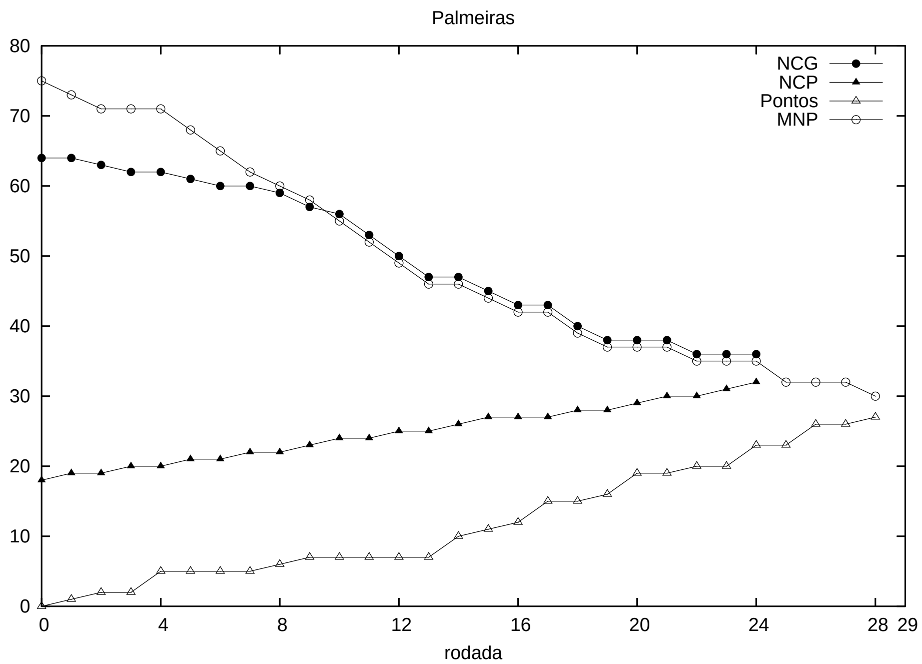

Figura 6.2: Palmeiras no Campeonato Brasileiro 2002.

6.5 Conclusões

Neste capítulo foi considerada uma aplicação de programação inteira para determinar a situação das equipes participantes de um campeonato em relação aos objetivos de classificação. Os dois modelos propostos determinam condições necessárias e suficientes para a classificação de uma determinada equipe nas primeiras k posições.

O sítio web do Projeto FutMax [48] continua accessível e atualizado. O projeto acompanhou todas as edições do Campeonato Brasileiro a partir de 2002. Várias mudanças nas regras do campeonato, assim como algumas penalidades impostas a determinadas equipes, tornaram necessárias algumas modificações nos modelos, que se mostraram muito flexíveis. Novos objetivos, como a classificação para a Taça Libertadores, e novos campeonatos, como as eliminatórias sul-americanas para a Copa do Mundo na Alemanha, foram e estão sendo acompanhados. Novas estatísticas de interesse do público foram disponibilizadas.

Como conseqüência do trabalho desenvolvido neste capítulo, foram publicados os artigos [50, 52]

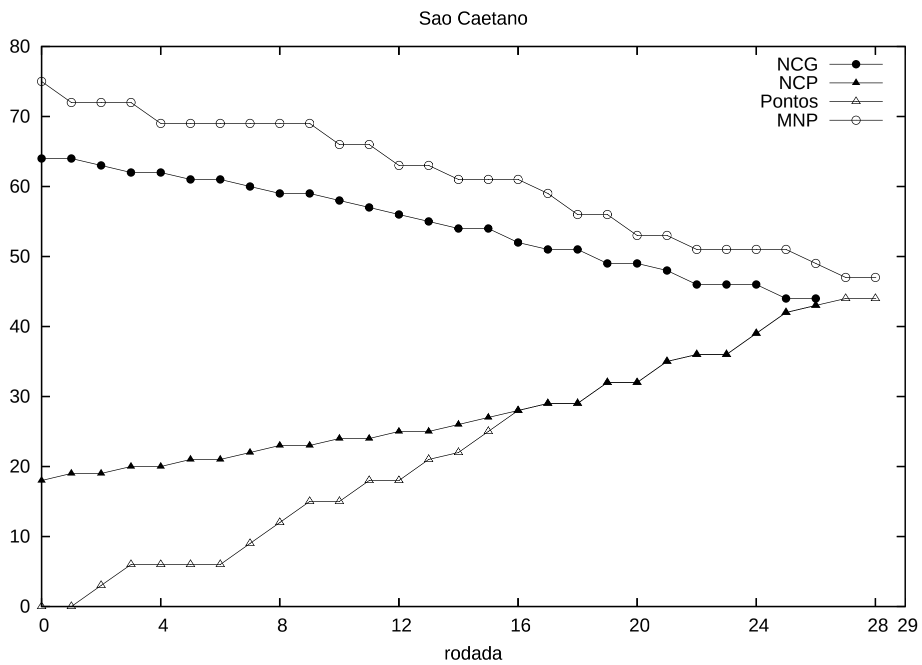

Figura 6.3: São Caetano no Campeonato Brasileiro 2002.