Aplicando o Golden Ball ao Problema da Diversidade Máxima

Carlos V. Dantas Araújo[1], Jailon W. B. Oliveira da Silva[1], Pablo L. Braga Soares[1]

[1]Universidade Federal do Ceará – Campus Russas
Avenida Felipe Santiago – Nº 411 62.900-624, Russas/CE, Brasil

{carlosvinicio, williambrunocc}@alu.ufc.br, pablo.soares@ufc.br

Abstract. *This work adapts the Golden Ball metaheuristic to the Maximum Diversity Problem in graphs, modifying the sports-inspired algorithm to select subsets of vertices that maximize the sum of their distances. Experiments with 80 instances demonstrated competitive performance, achieving solutions with deviations ranging from 0.06% to 0.95% compared to the best-known values, in addition to a considerably shorter execution time compared to the Scatter Search method, a reference in the literature. The algorithm excelled in geometric instances, obtaining solutions equivalent to the best-known values in 95% of cases, and exhibited consistent performance across the remaining instances.*

Resumo. *Este trabalho adapta a meta-heurística Golden Ball ao Problema da Diversidade Máxima em grafos, modificando o algoritmo inspirado em esportes para selecionar subconjuntos de vértices que maximizam a soma de suas distâncias. Experimentos com 80 instâncias demonstraram um desempenho competitivo, alcançando soluções com desvio de 0,06% a 0,95% em relação aos melhores valores conhecidos, além de um tempo de execução consideravelmente menor em comparação ao método Scatter Search, referência na literatura. O algoritmo se destacou em instâncias geométricas, obtendo soluções equivalentes aos melhores valores conhecidos em 95% dos casos, e apresentou desempenho consistente nas demais instâncias.*

1. Introdução

Na área da Teoria dos Grafos, diversos problemas combinatórios são classificados como NP-difíceis, ou seja, problemas cuja solução exata é inviável em grandes instâncias devido ao crescimento exponencial do tempo computacional necessário [Garey and Johnson 1979]. Dentre esses problemas, destaca-se o Problema da Diversidade Máxima, do inglês Maximum Diversity Problem (MDP), que visa selecionar subconjuntos de vértices em um grafo de forma a maximizar a soma das distâncias entre os vértices escolhidos [Gonzalez et al. 2020].

O MDP é formalmente definido no contexto de grafos da seguinte maneira: dado um grafo $G = (V, E)$, onde V representa o conjunto de vértices e E o conjunto de arestas, o objetivo é selecionar um subconjunto $S \subseteq V$ que maximize a soma das distâncias entre os vértices selecionados. Sendo um problema NP-difícil, a solução exata para instâncias de grande escala é impraticável, uma vez que o tempo computacional cresce exponencialmente com o aumento do tamanho do grafo. Esta restrição impõe a necessidade de se adotar abordagens aproximadas, como heurísticas e meta-heurísticas, que resolvem o problema dentro de um tempo aceitável, embora com uma perda de precisão [Blum and Roli 2003].

Embora o uso da meta-heurística Golden Ball em problemas combinatórios, como o Problema do Roteamento de Veículos [Worawattawechai et al. 2022] e o Problema do Roteamento de Veículos Capacitados [Ruttanateerawichien et al. 2014], já esteja presente na literatura, sua aplicação ao MDP ainda é pouco explorada. Este trabalho busca preencher essa lacuna ao investigar como essa meta-heurística pode ser empregada na resolução de instâncias do MDP, visando melhorar a eficiência computacional e a qualidade das soluções.

2. Definição Formal e Abordagem Adotada

O Problema da Diversidade Máxima é uma formulação combinatória que busca selecionar um subconjunto de vértices de modo a maximizar a diversidade entre eles. Dado um grafo $G = (V, E)$, onde V é o conjunto de vértices e E é o conjunto de arestas, o objetivo do MDP é encontrar um subconjunto $S \subseteq V$ de tamanho m que maximize a soma das diferenças entre os vértices em S. Essas diferenças são geralmente medidas pelos pesos das arestas, de modo a maximizar a separação entre os vértices selecionados [Gonzalez et al. 2020, Duarte et al. 2015, Carrabs et al. 2022].

Formalmente, o problema pode ser expresso da seguinte forma:

$$\max_{S \subseteq V, |S| = m} \sum_{v_i, v_j \in S} d(v_i, v_j) \tag{1}$$

onde $d(v_i, v_j)$ representa a medida de diferença entre os vértices v_i e v_j. O MDP é um problema NP-difícil [Ghosh 1996], o que significa que não há um algoritmo eficiente conhecido para resolvê-lo de forma exata em tempo polinomial, especialmente em grafos de grande escala. Por isso, técnicas heurísticas e meta-heurísticas são frequentemente empregadas para encontrar soluções aproximadas.

Para abordar o MDP, adaptamos a meta-heurística Golden Ball, um algoritmo multipopulacional inspirado em competições esportivas [Osaba et al. 2014]. Sua estrutura formal compreende sete componentes:

$$GB = \langle P, \mathcal{F}, \mathcal{C}, \mathcal{L}, \mathcal{T}, \mathcal{R}, \tau \rangle \tag{2}$$

onde:

- $P = \{S_1, \ldots, S_m\}$ representa a população de m soluções candidatas;
- $\mathcal{F}$ é a função que avalia a qualidade de cada solução;
- $\mathcal{C}$ é o operador que divide as soluções em grupos competitivos;
- $\mathcal{L}$ implementa os procedimentos de busca local;
- $\mathcal{T}$ gerencia a transferência de componentes entre soluções;
- τ define os critérios de parada do algoritmo.

O algoritmo opera em ciclos iterativos compostos por cinco fases sequenciais:

1. **Avaliação**: Cada solução S_i na população P é avaliada através da função $\mathcal{F}(S_i)$, que no contexto do MDP calcula a soma das distâncias entre os vértices selecionados;

2. **Competição**: A população é dividida em dois grupos através do operador $\mathcal{C}$: as soluções elite P_e (20% melhores) e as não-elite P_n (80% restantes), criando uma hierarquia que irá guiar as fases seguintes;
3. **Melhoria Local**: Cada solução passa por um processo de refinamento através do operador $\mathcal{L}$, que testa substituições locais de vértices. Para uma solução S_i, são avaliadas trocas do tipo $\{v\} \leftrightarrow \{u\}$, onde $v \in S_i$ e $u \in V \setminus S_i$, mantendo apenas as alterações que aumentam o valor de $\mathcal{F}(S_i)$;
4. **Transferência**: Soluções do grupo P_n podem receber vértices promissores das soluções elite P_e através do operador $\mathcal{T}$. Especificamente, para cada $S_i \in P_n$, identificamos o vértice $v \in S_i$ com menor contribuição para a diversidade (minimizando $\sum_{w \in S_i} d(v, w)$) e o substituímos por um vértice de alguma $S_j \in P_e$;
5. **Renovação**: Soluções que não apresentam melhora após várias iterações são parcialmente reconstruídas pelo operador $\mathcal{R}$, que preserva os $\lceil k/2 \rceil$ vértices com maior contribuição individual e completa a solução com novos vértices selecionados aleatoriamente, seguindo estratégias comprovadamente eficazes em problemas de diversidade [Aringhieri et al. 2015];

Para garantir eficiência computacional, três otimizações foram implementadas:

- Um mecanismo de **cache** que armazena os valores de $\mathcal{F}(S_i)$ já calculados, evitando reprocessamento desnecessário;
- O cálculo de **atualização incremental** que permite avaliar trocas de vértices de forma eficiente através da expressão:

$$\Delta \mathcal{F} = \sum_{w \in S_i} (d(u, w) - d(v, w)) \tag{3}$$

  onde v é o vértice sendo removido e u o vértice candidato a entrar na solução, uma técnica amplamente utilizada em problemas de diversidade [Martí et al. 2013];
- A **avaliação paralela** que permite processar diferentes soluções simultaneamente, reduzindo significativamente o tempo de execução.

O algoritmo encerra sua execução quando atinge um dos seguintes critérios de parada τ: (1) completar 100 iterações consecutivas sem melhoria na melhor solução encontrada, ou (2) atingir o limite máximo de 1200 segundos de processamento.

Esta adaptação do Golden Ball ao MDP mantém os princípios fundamentais da meta-heurística original enquanto introduz algumas especializações necessárias para lidar com as particularidades do Problema da Diversidade Máxima.

3. Resultados Computacionais

Os experimentos foram conduzidos em um sistema equipado com processador Intel® Core™ i5-12450H (2.00 GHz) e 16GB de memória RAM, utilizando implementações em Python. Neste estudo, n representa o número total de vértices no grafo e m o tamanho do subconjunto a ser selecionado para maximizar a diversidade. O algoritmo Golden Ball (GB) foi avaliado em 80 instâncias da MDPLIB, abrangendo três categorias principais: SOM-b (20 instâncias com $n \in \{100, 200, 300, 400, 500\}$ e m variando de 10 a 200), GKD-c (20 instâncias com $n = 500$ e $m = 50$) e MDG (40 instâncias divididas em MDG-a e MDG-b, todas com $n = 500$ e $m = 50$).

Para comparação, utilizamos como referência o algoritmo *Scatter Search with Memory Structures* (G_SS), desenvolvido por [Gallego et al. 2009] e reconhecido como o método baseado em população que obteve os melhores resultados conhecidos na literatura para o Problema da Diversidade Máxima. O G_SS foi executado com 2 horas por instância em seus experimentos originais e seus resultados possuem acesso público em https://grafo.etsii.urjc.es/optsicom.

Tabela 1. Comparação entre o Golden Ball (GB) e o Scatter Search (G_SS)

Categoria	**Desvio Médio (%)**	**Melhores Valores (%)**	**Tempo de Execução**
SOM-b	0.82	25.0	20 min
GKD-c	0.06	95.0	20 min
MDG-a	0.91	5.0	20 min
MDG-b	0.95	5.0	20 min
G_SS (referência)	0.00	100.0	120 min

O GB demonstrou bom desempenho com apenas um sexto do tempo de execução do método G_SS. Para as instâncias GKD-c, que utilizam distâncias euclidianas, o algoritmo obteve desvio médio de **0.06%** em relação aos melhores valores conhecidos, alcançando soluções equivalentes às de referência em 95% dos casos. Esse resultado sugere que a dinâmica de competição entre times do Golden Ball é particularmente eficaz em problemas com estrutura geométrica.

Nas instâncias SOM-b, o GB apresentou desvio médio de **0.82%**, com 25% das soluções atingindo os mesmos valores de referência. O desempenho foi ainda mais notável em instâncias menores ($m \leq 200$), onde os desvios não ultrapassaram 0.5%. Para as categorias MDG-a e MDG-b, tradicionalmente mais desafiadoras, os desvios médios foram de **0.91%** e **0.95%** respectivamente.

O algoritmo completou em média **45 iterações** por execução, com tempo médio de **25-30 segundos/iteração** para instâncias com $m \geq 100$. Em todos os casos, o GB manteve capacidade exploratória ativa, sem atingir o critério de parada por estagnação. O maior desvio individual observado foi de **0.98%** na instância MDG-b_15, enquanto em 60% das execuções os desvios ficaram abaixo de 0.5%.

- **Eficiência comprovada**: Soluções com desvios <1% em 20 minutos;
- **Precisão notável**: 95% de soluções equivalentes nas instâncias GKD-c;
- **Consistência**: Desempenho consistente em todos os grupos de instâncias.

Estes resultados demonstram que o Golden Ball estabelece uma nova perspectiva na relação tempo-desempenho para o Problema da Diversidade Máxima, oferecendo soluções promissoras em tempo de execução significativo.

4. Conclusão

O uso da meta-heurística Golden Ball para a resolução do Problema da Diversidade Máxima é uma contribuição inédita deste trabalho. Os resultados preliminares indicam um caminho promissor para a obtenção de melhores soluções para o MDP. Trabalhos futuros incluem aprimoramentos e técnicas híbridas, além de testes com maior tempo de execução.

Referências

Aringhieri, R., Bruglieri, M., Malucelli, F., and Nonato, M. (2015). Special issue on optimization in medicine and biology: Exact and heuristic approaches. *European Journal of Operational Research*, 244(3):647–648.

Blum, C. and Roli, A. (2003). Metaheuristics in combinatorial optimization: Overview and conceptual comparison. *ACM Computing Surveys (CSUR)*, 35(3):268–308.

Carrabs, F., Cerulli, R., and Toth, P. (2022). New approaches for solving the maximum diversity problem in large graphs. *Journal of Heuristics*, 28(3):231–256.

Duarte, A., Martí, R., Sánchez-Oro, J., and Glover, F. (2015). Metaheuristics for the maximum diversity problem. *European Journal of Operational Research*, 240(1):24–36.

Gallego, M., Duarte, A., Laguna, M., and Martí, R. (2009). Hybrid heuristics for the maximum diversity problem. *Computational Optimization and Applications*, 44(3):411–426.

Garey, M. R. and Johnson, D. S. (1979). *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W.H. Freeman and Company.

Ghosh, J. B. (1996). Computational aspects of the maximum diversity problem.

Gonzalez, J. A., Hernández-Pérez, H., and Lozano, M. (2020). A review on maximum diversity problems. *European Journal of Operational Research*, 287(2):401–420.

Martí, R., Duarte, A., and Laguna, M. (2013). Advanced scatter search for the max-min diversity problem. *INFORMS Journal on Computing*, 25(2):245–260.

Osaba, E., Diaz, F., and Onieva, E. (2014). Golden ball: A novel meta-heuristic to solve combinatorial optimization problems based on soccer concepts. *Applied Intelligence*, 41(1):145–166.

Ruttanateerawichien, K., Kurutach, W., and Pichpibul, T. (2014). An improved golden ball algorithm for the capacitated vehicle routing problem. In Pan, L., Păun, G., Pérez-Jiménez, M. J., and Song, T., editors, *Bio-Inspired Computing - Theories and Applications*, pages 341–356, Berlin, Heidelberg. Springer Berlin Heidelberg.

Worawattawechai, T., Intiyot, B., Jeenanunta, C., and Ferrell, W. G. (2022). A learning enhanced golden ball algorithm for the vehicle routing problem with backhauls and time windows. *Computers Industrial Engineering*, 168:108044.