BOLETIM APDIO

PUBLICAÇÃO SEMESTRAL | 66 | JUNHO '22

ENTREVISTA

JOSÉ FERNANDO GONÇALVES

IO EM AÇÃO

PROGRAMAÇÃO DE ROTAS NA ÚLTIMA MILHA O AMAZON LAST MILE ROUTING RESEARCH CHALLENGE

TÉCNICAS DE IO

OTIMIZAÇÃO SOB INCERTEZA DE DADOS

ARTIGO DE OPINIÃO

OS DESAFIOS DO SETOR DO RETALHO NA ERA DIGITAL

ÍNDICE

03 **ENTREVISTA**
JOSÉ FERNANDO GONÇALVES

04 **ARTIGO DE OPINIÃO**
OS DESAFIOS DO SETOR DO RETALHO NA ERA DIGITAL
Susana Relvas

08 **TÉCNICAS DE IO**
OTIMIZAÇÃO SOB INCERTEZA DE DADOS
Marco Silva
João Pedro Pedroso

11 **IO EM AÇÃO**
PROGRAMAÇÃO DE ROTAS NA ÚLTIMA MILHA
O AMAZON LAST MILE ROUTING RESEARCH CHALLENGEALIMENTAR
António Galrão Ramos
Manuel Pereira Lopes

15 **LUGAR AOS NOVOS**
(DIS)SIMILARIDADE EM PROCESSOS DE TOMADA DE DECISÃO COM MÚLTIPLOS CRITÉRIOS E CATEGORIAS NOMINAIS
Ana Sara Costa

19 **PT EM IO PELO MUNDO**
A IMPORTÂNCIA DA INVESTIGAÇÃO OPERACIONAL NUMA EMPRESA DE LOGÍSTICA PARA E-COMMERCE EM RÁPIDA EXPANSÃO PELA EUROPA
Bruno Vieira

20 **NOTÍCIAS DA APDIO**

ELSA SILVA
INESC TEC
Universidade do Minho

MARIA JOÃO SANTOS
INESC TEC
Universidade do Minho

SARA MARTINS CORREIA
Escola Superior de Tecnologia e Gestão
Politécnico do Porto
INESC TEC

EDITORIAL

A Investigação Operacional (IO) tem por função principal a otimização. Este Boletim será dedicado à aplicação da IO no e-commerce (ou comércio eletrónico). São muitos e variados os desafios que emergem no e-commerce para os quais a IO pode e tem contribuído com soluções eficientes.

A Entrevista a José Fernando Gonçalves, *Principal Research Scientist* na Amazon, apresenta alguns dos problemas em que a IO é utilizada no maior retalhista online do mundo, que vão desde o problema de roteamento, relocalização de reboques, gestão de cacifos, gestão de parques de estacionamento até ao escalonamento da manutenção de aviões. Refere ainda quais são, na sua opinião, os principais desafios que o e-commerce enfrentará e como poderá a IO contribuir para a sua resolução.

No Artigo de Opinião, Susana Relvas aborda os principais desafios do setor do retalho na era digital. Define o que é o retalho e quais os diferentes atores no sistema que compõe este setor. Apresenta uma nota histórica para o surgimento e desenvolvimento do e-commerce e quais as áreas de investigação com maior número de publicações no tópico *retail operations*. Uma breve visão do estado da arte é apresentada assim como os principais desafios para a área do retalho.

Na rúbrica Técnicas de IO, Marco Silva e João Pedro Pedroso, abordam a otimização sob incerteza de dados. Apresentam-se as duas categorias de modelação do problema considerando a incerteza e motiva-se a sua utilização. Posteriormente apresenta-se uma descrição formal do problema e discutem-se os diferentes métodos de solução que podem ser adotados. A rúbrica termina com exemplos de aplicação em particular no e-commerce.

António Galrão Ramos e Manuel Pereira Lopes, descrevem a competição científica Amazon Last Mile Routing Research Challenge na qual participaram, tendo alcançado o 4.ª lugar num total de 229 equipas. Para além da descrição do problema descrevem também a abordagem adotada para o resolver e qual o método de avaliação aplicado pelos organizadores da competição.

Na rúbrica Lugar os Novos, Ana Sara Costa descreve o método *CATegorization by Similarity-Dissimilarity* proposto por si, assim como a sua aplicação no Exército Português para o recrutamento e seleção dos militares das tropas especiais do exército.

Bruno Vieira, Analista de Investigação Operacional na Paack em Barcelona, partilha connosco na rúbrica PT em IO pelo mundo, o seu percurso académico e profissional na IO e os desafios recentes que tem superado na *scale-up* na área de logística para e-commerce.

O Boletim termina com o habitual espaço reservado às notícias relevantes para a comunidade de IO.

Gostaríamos de agradecer à Comissão Editorial anterior constituída pela Eliana Costa e Silva e pelo Rui Borges Lopes pelo trabalho extraordinário realizado nos últimos quatro anos. Apresentaram boletins de excelente qualidade com temas de extrema relevância para a comunidade de IO, que ficarão para sempre na história do Boletim da APDIO. Assumimos o compromisso de, nesta e nas próximas edições, nos empenharmos para manter o legado de elevada qualidade das comissões editoriais anteriores na prossecução de temas do interesse da comunidade de IO.

ENTREVISTA

Having an impressive academic curriculum in the OR field with solutions for different problems the area you are most well-known is Biased Random-key Genetic Algorithms. Can you share with us the history of BRKGA and what makes it so useful and performant in solving so many different types of problems?

Biased random-key genetic algorithms are a variant of random-key genetic algorithms introduced by Bean (ORSA Journal Computing 6:154–160, 1994) and includes the following two innovations:

- **A new mating process** (crossover) where one of the parents used is biased to be of higher fitness, i.e., is an elite, or high-quality, solution;
- **A new way of envisioning random-key genetic algorithms as a general framework** (or metaheuristic) that clearly separates the problem-independent part of the architecture, the evolutionary engine, from the problem-specific part, the decoder, which associates to a vector of random keys a solution of the combinatorial optimization problem and a fitness value (measure of quality).

Experiments run to construct runtime distributions, for several types of problems, demonstrated that the biased variant (using the new mating process) is faster and obtains better results than Bean's algorithm. The new BRKGA framework makes clear the following characteristics of biased random-key algorithm:

- The evolutionary engine is always the same, thus allowing the reuse of the same software implementation;
- Any improvements made to the evolutionary engine are automatically incorporated in any problem previously solved without causing incompatibilities;
- The algorithm designer needs only to concentrate on the design of the decoder;
- Several decoders can be used for the same problem;
- Any constructive heuristic for a problem can be used as a decoder and can easily take advantage of the evolutionary engine by randomizing some of its parameters;
- The use of knowledge specific to a problem on hand can easily be incorporated into the decoder (e.g. a job-shop problem decoder should only produce active schedules);
- Any type of local search improvement can be included in the decoder.

In 2017 you moved to Amazon, the world largest online retailer, as Principal Research Scientist. Operations Research must be used in many areas of Amazon. Can you share your experience with OR in Amazon and what problems you have been working?

Amazon is a paradise for OR people. Amazon has problems that cover almost all areas of OR. Below I list several areas of OR where I have worked for Amazon:

Routing – This is the main area in terms of importance in the logistic process of Amazon. I worked on generating routes between internal facilities using VRPs with one load per leg and time windows and generating routes for Amazon FLEX drivers (UBER like drivers) using a VRP with pickups and deliveries and time windows. Both problems are very difficult because of the very large number of orders (>100000). I used BRKGA to generate routes.

Relocation of Trailers – The products are transported between Amazon facilities using trailers. Empty trailers have to return to pick up more products but this is a task that does add value so we need to minimize its cost. I used a MIP which I reformulated in order to convert it to an LP with integrality property.

Yield management of lockers – Amazon uses lockers to deliver orders. However, some customers take a long time to pick up the order and therefore prevent the usage of the locker by other customers. I developed a Simulation-Optimization approach that determines the policy used to determine if a customer should be allowed to use lockers to pick a particular order if the lockers are constrained.

Simulation of parking yards – The objective here was to determine what were the factors (rate of arrival, parking time, etc.) that would cause congestion in the parking yard and which ones were the most important to control.

Scheduling of aircraft engine maintenance – I created a MIP model to determine how many engines we should buy or lease and when the engines should go maintenance. The MIP could be solved in a very short time using a commercial software like XPRESS from FICO.

JOSÉ FERNANDO GONÇALVES

Principal Research Scientist at Amazon.com
jfgoncal@gmail.com

OR is also applied at Amazon in areas like:

Inventory management, Location of facilities, Scheduling of trips, design of auctions to get services, and forecasting of demand, design of auctions to get drivers or other types of services, attendance of drivers, price surge, etc.

In your opinion what are the main future challenges of e-commerce and how can OR contribute to solve them?

The main future challenges of e-commerce will be the **CO2 footprint** produced by the deliver process and the use of **environmentally friendly materials** in the packages. Amazon is already trying to solve both problems. OR is being used to solve the CO2 footprint by combining **electric vehicle (EVs)** with **swarm of drones (SOW)**. The EV-SOW solution is being studied using OR approaches similar to the ones used for making routes (VRPTW, VRPPD, etc.) and the ones used for location problems.

OS DESAFIOS DO SETOR DO RETALHO NA ERA DIGITAL

SUSANA RELVAS

Departamento de Engenharia e Gestão
Centro de Estudos de Gestão
Instituto Superior Técnico
susana.relvas@tecnico.ulisboa.pt

RETALHO – A INDÚSTRIA E UM TRONO EM DISPUTA

Quando se menciona a palavra "retalho", imediatamente a associamos ao processo de compra de bens, em loja ou no canal online. No entanto, a indústria de retalho envolve mais atores do que aqueles que são visíveis ao consumidor. Consiste num sistema que envolve consumidores, retalhistas, distribuidores, produtores, fornecedores ou prestadores de serviços (figura 1). Envolve ainda transações não só com consumidores (*business to consumer* ou B2C) mas também entre diferentes negócios que se constituem como clientes entre si (*business to business* ou B2B) [1]. Para que estas transações e a venda efetiva de bens ou serviços possam ocorrer e satisfazer o cliente ou consumidor, estes atores desenvolvem um conjunto de atividades associadas à indústria de retalho. Estas atividades incluem a gestão da cadeia de abastecimento e das relações entre os atores envolvidos, gestão de inventário, logística e distribuição, gestão do consumidor, merchandising, operações de loja ou gestão de preço [2].

Claramente, a gestão (e a investigação) destas atividades recai na interface entre o marketing e a gestão de operações, e nem sempre é abordada em harmonia. O marketing tem o desafio de saber o que o mercado procura, enquanto as operações verificam como é que a organização consegue satisfazer essas expectativas. O marketing prefere ter a prateleira cheia enquanto as operações preferem otimizar o inventário e reduzir desperdício. Em virtude de aumentar as vendas e satisfazer o cliente, o marketing quer aceitar qualquer encomenda independentemente da existência de capacidade operacional para a satisfazer. Ou o marketing prefere variedade enquanto as operações têm tendência a estandardizar os produtos, processos e atividades.

Como diria Tywin Lannister na famosa série "Game of Thrones", "quando os soldados não têm disciplina, a culpa reside no seu comandante". Em retalho, o objetivo deve ser claro e único e é nessa premissa que deve ser abordado o estado atual do setor, a sua gestão e a investigação associada.

OS DESAFIOS ATUAIS DO SETOR DO RETALHO

O comércio eletrónico (do inglês *electronic commerce* ou *e-commerce*) surgiu, ele próprio, da necessidade de "encomendar" bens na Alemanha pós II Guerra Mundial nos finais dos anos 40. Numa Berlim dividida pelas intervenções dos Aliados e da União Soviética, surgiu um sistema de encomendas via telex. Mais tarde, na década de 70, surgiu o primeiro *standard* de comunicação computador-para-computador através de um sistema de *Electronic Data Interchange*. Outros eventos mais ou menos disruptivos surgiram na história do retalho, como por exemplo o aparecimento de Operadores Logísticos com a diminuição da regulação do setor, o crescimento de grandes cadeias de retalho e a transferência da influência dos produtores para os retalhistas, o aparecimento de hipermercados e outros tipos de formatos comerciais, novas estratégias comerciais, a miniaturização dos produtos, primeiro a globalização e agora o regresso ao "local", entre outros. A constante capacidade de adaptação do setor continua diariamente,

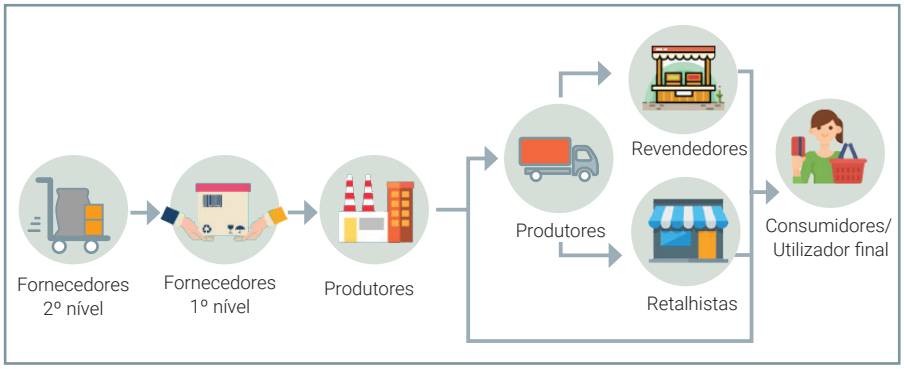


Fig 1 - Atores na indústria de retalho

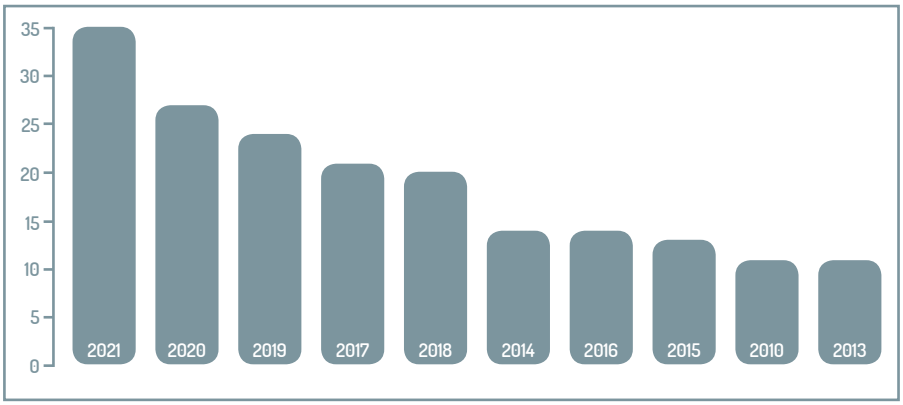


Fig 2 - Anos com maior número de publicações em "retail operations" (no tópico), segundo a Web of Science (consultado a 13 de maio de 2022)

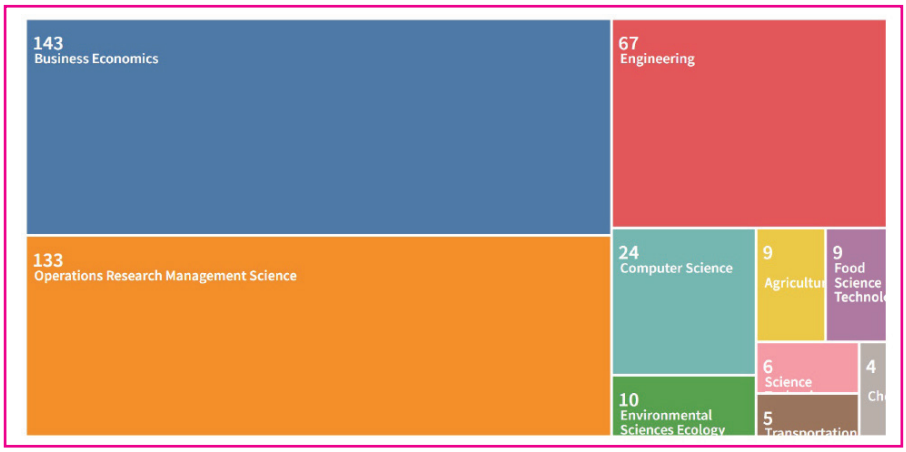


Figura 3 – Áreas de investigação maior número de publicações em "retail operations" (no tópico), segundo a Web of Science (consultado a 13 de maio de 2022)

mesmo quando desafiada por eventos extremos como a pandemia provocada pela Covid-19 ou a guerra na Ucrânia. Segundo Mike George, chairman da National Retail Federation, "o consumidor sabe o que os retalhistas e atores envolvidos no setor são capazes de fazer e já têm expectativas muito elevadas face a novos desafios" [3].

O consumidor em 2022 procura ter de novo mais experiências fora de casa e ter experiências semelhantes independentemente do canal de compra. Há ainda uma tendência crescente para o consumidor se identificar com retalhistas e organizações que têm mais apetência por práticas que os alinhem positivamente com os critérios ESG (do inglês *Environmental, Social and Governance*) [4], estando mesmo disponível a pagar um valor extra para se alinhar com objetivos sustentáveis [5]. Nunca foi tão relevante conhecer os estilos de vida, os valores e as motivações dos consumidores, em especial quando a geração "alpha" (nascida no novo milénio, totalmente enquadrada na era tecnológica) começa a ser um consumidor direto [3], [5]. O consumidor aprende sobre, procura e compra os produtos cada vez mais de forma híbrida. Também tem maior tendência a experimentar e adotar o uso de tecnologia [5].

Do lado do retalhista, as principais estratégias a adotar pelos executivos incluem aumentar as suas capacidades digitais, adap-

"A CONSTANTE CAPACIDADE DE ADAPTAÇÃO DO SETOR CONTINUA DIARIAMENTE, MESMO QUANDO DESAFIADA POR EVENTOS EXTREMOS COMO A PANDEMIA PROVOCADA PELA COVID-19 OU A GUERRA NA UCRÂNIA"

"CLARAMENTE, A GESTÃO (E A INVESTIGAÇÃO) DESTAS ATIVIDADES RECAI NA INTERFACE ENTRE O MARKETING E A GESTÃO DE OPERAÇÕES"

tar a força de trabalho para estar pronta para o futuro, mas enfrentando falta de talento nas áreas de tecnologias de informação e analítica empresarial, readaptar as lojas ao conceito omnicanal, aumentar a segurança e privacidade de dados, ou incorporar práticas alinhadas com os critérios ESG. Os executivos acreditam ainda que o consumidor vai valorizar a disponibilidade de inventário em detrimento da lealdade ao retalhista, reagindo às falhas recentes das cadeias de abastecimento e impondo uma gestão resiliente [4], pois o próprio consumidor também se tornou resiliente [5].

Por fim, do ponto de vista do desenvolvimento económico, a União Europeia define a Indústria 5.0 como o paradigma a adotar pela indústria em geral. De modo a se manter próspera, deve equilibrar os objetivos de eficiência e produtividade com o seu contributo para a sociedade. Este contexto consiste em complementar as diretrizes da Indústria 4.0 com a transição para uma indústria focada no humano, mais sustentável e mais resiliente [6].

Cruzando o consumidor, o retalhista e o contexto da União Europeia, a evolução do retalho centra-se nas pessoas e não nas corporações, o que leva a uma individualização da experiência do consumidor. Se o consumidor precisar de um produto, ele vai lá estar disponível. E será o produto certo para si.

OPERAÇÕES EM RETALHO – BREVE VISÃO DO ESTADO DA ARTE

A já referida fronteira entre o marketing e a gestão de operações está subjacente à temática das operações em retalho. Alinhada com o dinamismo do setor do retalho, a comunidade científica tem vindo a abordar o tema das operações em retalho com mais relevância. Na figura 2 identificam-se os anos com maior número de publicações com o tópico de "retail operations", segundo a Web of Science. O número total de publicações é de 268, tendo sido publicadas 190 nos 10 anos mais relevantes. Os anos com maior número de publicações situam-se de 2010 em diante. O ano de 2022, apesar de ainda não constar na figura, já é o 11º ano com maior número de publicações relacionadas com "retail operations".

Fazendo a análise a estas publicações do ponto de vista das áreas de investigação onde foram publicadas, o resultado obtido encontra-se na figura 3. Verifica-se um equilíbrio no topo entre "Business Economics" e "Operations Research & Management Science". Este equilíbrio corrobora a tese da fronteira entre o Marketing e a Gestão de Operações, mas traz ainda a relevância da Investigação Operacional para a investigação em "Retail Operations".

De notar que o número de publicações com o tópico "retalho" (*retail*) ascendem a mais de 50 000, em que mais de 1200 são artigos de revisão. Neste conjunto de publicações, a preponderância das áreas de investigação de "Business", Economics" e "Management" é muito elevada. Há ainda relevância de publicações relacionadas com sustentabilidade em retalho, muito associada aos setores alimentar e vestuário.

Focando em artigos de revisão relevantes para a temática de "retail operations" retirados do conjunto de artigos de revisão em "retail", enumeram-se algumas das principais mensagens de publicações recentes e que estão alinhadas com os desafios identificados neste artigo:

- No trabalho de Melacini e co-autores [7] é clara a **necessidade de investigar como devem ser desenvolvidas as redes de retalho**, nomeadamente em termos de localizações e papel de cada localização, como planear a variedade de produto oferecida nos diferentes canais e a integração no contexto omnicanal, e como articular as operações logísticas em geral e em particular que papel logístico devem ter as lojas. Também é discutida a problemática da gestão de capacidade operacional e as dificuldades que advêm da implementação de um formato omnicanal do ponto de vista das operações.
- A constante **evolução do tema de "supply chain analytics" em detrimento de "supply chain optimization"** e a oportunidade de utilização de métodos baseados em dados aplicados à temática de retalho foram analisados no trabalho de Qi e co-autores [8]. Foram detalhados modelos baseados em dados aplicados a problemas de variedade de produtos (*assortment*), problemas de satisfação de encomendas ou problemas de gestão de inventário. No entanto, os autores referem que quer nestes problemas quer noutros relacionados com operações em retalho existem oportunidades de investigação, nomeadamente criar conhecimento sobre a sua adequação e a sua estruturação.

"SE O CONSUMIDOR PRECISAR DE UM PRODUTO, ELE VAI LÁ ESTAR DISPONÍVEL. E SERÁ O PRODUTO CERTO PARA SI"

O FUTURO DO RETALHO SERÁ SEM DÚVIDA MUITO EXCITANTE

• Já no trabalho de Cai e Lo [9], focado na temática omnicanal, é relevada a **necessidade de examinar e sistematizar o papel da tecnologia** na adoção e gestão deste formato. Adicionalmente, identificam que as preferências do consumidor podem ser relativas ao produto ou ao serviço prestado. São indicadas três direções de trabalho: i) a redução ou eliminação das discrepâncias entre o formato online e offline, ii) o desenho do sistema produto-serviço e iii) a incorporação das preocupações de sustentabilidade no setor.

• O trabalho de Ahsan e Rahman [10] é muito recente e concentra-se na **problemática da devolução de produtos**, nomeadamente daqueles adquiridos em formatos online. São identificados cinco tópicos associados ao problema, nomeadamente a política de devolução, o processo de devolução, o serviço de devolução, custo e risco associado. Identificou-se que em geral o estudo dos recursos associados ao problema de devolução não é abordado, a saber: recursos humanos, recursos tecnológicos e recursos de tecnologias de informação. É ainda crítico determinar o comportamento do consumidor associado à devolução no contexto de encomendas internacionais ou via omnicanal e ainda a sustentabilidade ambiental ou social.

Como resultado de trabalho em curso no grupo de investigação da autora [11], foram identificadas de 17 decisões associadas a operações em retalho, deduzidas e homogeneizadas a partir da literatura: seleção de localizações, seleção de canais de distribuição, desenho da rede de distribuição, seleção de fornecedores, gestão da logística inversa, reposição de inventário em prateleira, operações de *backroom*, entrega, operações de *checkout*, previsão da procura, satisfação de encomenda, gestão de recursos humanos, gestão de inventário, variedade de produtos, planeamento de rotas, promoções e disposição de produtos. As decisões com mais publicações associadas são gestão de inventário e seleção de canais de distribuição. No que diz respeito a métodos de apoio à decisão utilizados, estes dividem-se entre otimização exata, uso de meta-heurísticas, simulação ou outros. Corrobora-se que a utilização de métodos baseados em dados é reduzida, apesar das oportunidades de trabalho identificadas por [8].

CONCLUSÃO

"*Amazon strives to be Earth's most customer-centric company, Earth's best employer, and Earth's safest place to work*". A mensagem incluída na visão do maior retalhista de e-commerce do mundo é clara na relevância do cliente e das pessoas. Apesar de no estado da arte haver muita referência ao consumidor, a clara identificação do que é uma operação de retalho centrada no cliente (*customer-centric*) e que decisões devem ser tomadas para o alcançar é um tópico pouco estudado [7]. Mas em retalho há pessoas também do lado da operação cujo papel no sucesso do retalho é crucial. Adiciona-se a necessidade de estruturar operações sustentáveis e resilientes, num contexto cada vez mais tecnológico. O futuro do retalho será sem dúvida muito excitante [2] e dinâmico e depende de cada um de nós.

"APESAR DE NO ESTADO DA ARTE HAVER MUITA REFERÊNCIA AO CONSUMIDOR, A CLARA IDENTIFICAÇÃO DO QUE É UMA OPERAÇÃO DE RETALHO CENTRADA NO CLIENTE (CUSTOMER-CENTRIC) E QUE DECISÕES DEVEM SER TOMADAS PARA O ALCANÇAR É UM TÓPICO POUCO ESTUDADO [7]. MAS EM RETALHO HÁ PESSOAS TAMBÉM DO LADO DA OPERAÇÃO CUJO PAPEL NO SUCESSO DO RETALHO É CRUCIAL"

REFERÊNCIAS

[1] *Ayers, James B.; Odegaard, Mary Ann (2018) Retail Supply Chain Management, CRC Press, Taylor & Francis Group, 2nd edition*

[2] *Caro, Felipe; Kök, A. Gürhan; Martínez-de-Albéniz, Victor (2020) The Future of Retail Operations. Manufacturing & Service Operations Management 22(1):47-58. https://doi.org/10.1287/msom.2019.0824*

[3] *Rethink Retail (2022) Acceleration Mode: Trends That Will Define 2022*

[4] *Deloitte (2022) 2022 retail industry outlook – The pandemic creates opportunities for the great retail reset, disponível em: https://www2.deloitte.com/content/dam/Deloitte/au/Documents/consumer-business/deloitte-au-cb-2022-retail-industry-outlook.pdf (consultado em maio de 2022)*

[5] *IBM (2022) Consumers want it all: Hybrid shopping, sustainability, and purpose-driven brands, disponível em https://www.ibm.com/downloads/cas/YZYLMLEV (consultado em maio de 2022)*

[6] *European Comission (ND) Industry 5.0, disponível em https://ec.europa.eu/info/research-and-innovation/research-area/industrial-research-and-innovation/industry-50_en (consultado em maio de 2022)*

[7] *Melacini, M., Perotti, S., Rasini, M. and Tappia, E. (2018), "E-fulfilment and distribution in omni-channel retailing: a systematic literature review", International Journal of Physical Distribution & Logistics Management, Vol. 48 No. 4, pp. 391-414. https://doi.org/10.1108/IJPDLM-02-2017-0101*

[8] *Qi, M; Mak, H-Y; Shen Z-JM (2020) Data-driven research in retail operations—A review, Naval Research Logistics, 67:595–616, https://doi.org/10.1002/nav.21949*

[9] *Cai, YJ; Lo, CK (2020) Omni-channel management in the new retailing era: A systematic review and future research agenda. International Journal of Production Economics 229:107729, https://doi.org/10.1016/j.ijpe.2020.107729*

[10] *Ahsan, K. and Rahman, S. (2022), "A systematic review of e-tail product returns and an agenda for future research", Industrial Management & Data Systems, Vol. 122 No. 1, pp. 137-166. https://doi.org/10.1108/IMDS-05-2021-0312*

[11] *del Zotti, Anna; Relvas, Susana (2022) Towards a customer-centricity framework for retail operations based on OR/OM Research, Working Paper*

OTIMIZAÇÃO SOB INCERTEZA DE DADOS

MARCO SILVA

INESC TEC
marco.c.silva@inesctec.pt

JOÃO PEDRO PEDROSO

INESC TEC e Faculdade de Ciências,
Universidade do Porto
jpp@fc.up.pt

A incerteza de dados em problemas de otimização tem sido tópico de interesse crescente. Desde o trabalho seminal de "Programação linear sob incerteza" de Dantzig em 1955 [6], onde foi introduzido o conceito de otimização estocástica, inúmeras abordagens foram desenvolvidas.

Com o aumento da maturidade dos métodos de resolução associados, a otimização sob incerteza de dados emergiu como o alicerce algorítmico para grande parte da estatística moderna e de *machine learning*. Suas aplicações podem ainda ser encontradas em um amplo espectro de problemas, incluindo planejamento financeiro, geração de eletricidade, gerenciamento da cadeia de suprimentos, mitigação das mudanças climáticas e controle da poluição, entre muitos outros.

MODELAGEM DA INCERTEZA

Uma dificuldade chave neste contexto é lidar com um grande espaço de incertezas, que frequentemente leva a modelos complexos e em grande escala. Por outro lado, incorporar o que se sabe sobre a incerteza nos parâmetros de entrada e variáveis utilizadas podem ajudar a quantificar a incerteza nas previsões resultantes.

Na medida em que probabilidades podem ser atribuídas às entradas, algumas dessas incertezas podem ser incorporadas. Esses modelos são assim chamados de modelos probabilísticos ou estocásticos.

A distribuição de tais incertezas pode estar disponível a partir de dados estatísticos, ou então pode ser somente um palpite — probabilidades "subjetivas". De qualquer forma, o tratamento matemático a ser dado é o mesmo. Uma ferramenta comum é trabalhar com cenários, que são representações particulares de como o futuro pode se desdobrar. O desafio, então, é como fazer bom uso dos cenários para com uma decisão eficaz.

Uma abordagem natural, mas falha, é resolver somente para cada cenário que é gerado. Esses problemas são eles próprios determinísticos em essência. Embora cada um produza uma prescrição do que deve ser feito no momento presente, não há orientação teórica sobre o compromisso entre essas prescrições e o que realmente deveria ser adotado, mesmo quando probabilidades podem ser atribuídos aos cenários individuais.

MODELAGEM DA DECISÃO

Por outro lado, problemas de otimização geralmente aparecem em um contexto temporal. A interação entre incerteza e tempo é importante para qualquer processo de tomada de decisão relacionado.

A literatura de otimização sob incerteza divide estes problemas em categorias como "esperar e ver (*wait and see*)" e "aqui e agora (*here and now*)". Em "esperar para ver" esperamos até que seja feita uma observação sobre os elementos aleatórios, e então resolvemos o problema, neste caso, determinístico.

Para a tomada de decisão do presente, as decisões devem ser feitas "aqui e agora". O problema do "aqui e agora" envolve a otimização sobre alguma medida probabilística - geralmente o valor esperado.

Uma abordagem comum é modelar o problema com o conceito de recurso. Neste caso, as variáveis de decisão do problema de otimização sob incerteza são divididos em dois conjuntos. Variáveis de primeiro estágio, são aquelas que devem ser decididas antes da concretização efetiva dos parâmetros incertos. São decisões do "aqui e agora". Subsequentemente, uma vez que os eventos aleatórios tenham se apresentado, melhorias no projeto ou na política operacional podem ser feitas selecionando-se, a um certo custo, os valores do segundo estágio, ou variáveis de recurso. São decisões do tipo "esperar para ver".

VALE A PENA?

Em muitos casos a abordagem de otimização estocástica é evitada porque ela parece complicar situações já complexas.

Uma questão natural é se perguntar se vale a pena lidar com a incerteza. A resposta teórica para isso é dada por dois indicadores: o valor esperado da informação perfeita (EVPI) e o valor da solução estocástica (VSS) [4].

O EVPI mede o quanto melhor um tomador de decisão poderia fazer se soubesse com certeza qual estado de natureza ocorreria.

Em situações em que não se pode reunir mais informações sobre o futuro, no entanto, pode ser mais pertinente que os tomadores de decisão saibam quão bem soluções de um modelo determinístico se compor-

tam em relação às soluções de programas estocásticos mais complicados. Assim, o VSS supõe que, em vez de resolver o problema estocástico, o decisor prefira resolver um problema de Valor Esperado, que consiste em substituir todas as variáveis aleatórias do problema estocástico pelos respetivos valores esperados. A diferença entre tomar o valor esperado do parâmetro incerto como a solução, como quando comparado ao uso de análise estocástica é definido como o VSS. Um método comum na literatura tem sido desenvolver limites fáceis de calcular para esses indicadores que possam orientar os profissionais a decidir a melhor maneira de lidar com a incerteza envolvida em seus problemas.

FORMULAÇÃO DO PROBLEMA

Podemos agora introduzir formalmente a formulação que tratamos neste contexto. Seja $x \in X \subseteq \mathbb{R}^n$ um vetor de decisão. Em um espaço mensurável $(\Xi, \mathcal{F})$, seja um vetor randômico $\xi \in \Xi \subseteq \mathbb{R}^d$, uma função de custo randômica $h(x,\xi) : X \times \Xi \to \mathbb{R}$, e um vetor de funções randômicas $g(x,\xi) : X \times \Xi \to \mathbb{R}^m$, i.e., $g(x,.) := [g_1(x,.), \ldots, g_m(x,.)]^T$. Dada esta definição, um problema genérico de otimização estocástica tem a forma

$$\inf_{x \in X} \{h(x) | \mathbb{E}_P[g(x,\xi)] \leq 0\} \qquad \text{(GSO)}$$

onde P designa a medida de probabilidade (conhecida) em $(\Xi, \mathcal{F})$ e $\mathcal{R}_P$: $\mathbb{Z} \to \mathbb{R}$ designa um funcional em P, onde $\mathbb{Z}$ é um espaço linear de funções mensuráveis em $(\Xi, \mathcal{F})$. O funcional $\mathcal{R}_P$ quantifica a incerteza inerente ao resultado das decisões, para uma dada medida de probabilidade P. Esta configuração representa uma ampla gama de problemas em estatística, otimização, e controle, tais como em problemas de regressão e classificação e processos de decisão de Markov.

Como um caso especial de (GSO), temos os problemas clássicos de otimização estocástica:

$$\inf_{x \in X} \mathbb{E}_P[h(x,\xi)], \qquad \text{(SO)}$$

e

$$\inf_{x \in X} \{h(x) | \mathbb{E}_P[g(x,\xi)] \leq 0\} \qquad \text{(CC)}$$

onde $\mathcal{R}_P[.]$ é tomado como o funcional de valor esperado $\mathbb{E}_P$. A formulação (CC) assume a forma de restrições probabilísticas (*chance constraints*). Os problemas (SO) e (CC) requerem o conhecimento da medida P [4]. Um problema de otimização robusto é definido como

$$\inf_{x \in X} \sup_{\xi \in \Xi} \{h(x,\xi) | \sup_{\xi \in \Xi} g(x,\xi) \leq 0\}, \qquad \text{(RO)}$$

O problema (RO) ignora todo o conhecimento distribucional de ξ, com exceção de seu suporte. O modelo minimiza o custo de pior caso dentro de um conjunto de incerteza [3]. Uma versão ambígua de (GSO) é formulada como

$$\inf_{x \in X} \sup_{P \in \mathcal{P}} \{\mathcal{R}_P[h(x,\xi)] | \sup_{P \in \mathcal{P}} \mathcal{R}_P[g(x,\xi)] \leq 0\}. \qquad \text{(DRO)}$$

Aqui, P denota um conjunto ambíguo (ambiguity set) de medidas de probabilidade, i.e., uma família of medidas consistente com o conhecimento prévio da incerteza. O problema (DRO) encontra a decisão que minimiza o pior caso do funcional R do custo entre todas as medidas de probabilidade do conjunto ambíguo desde que cada pior caso componente do funcional R da função $g(\ )$ seja não-positivo.

Esta abordagem é referida como otimização distribucional robusta (*distributionally robust optimization*), sob a sigla DRO. A abordagem (DRO) assume apenas um conhecimento par-

"UMA QUESTÃO NATURAL É SE PERGUNTAR SE VALE A PENA LIDAR COM A INCERTEZA. A RESPOSTA TEÓRICA PARA ISSO É DADA POR DOIS INDICADORES: O VALOR ESPERADO DA INFORMAÇÃO PERFEITA (EVPI) E O VALOR DA SOLUÇÃO ESTOCÁSTICA (VSS)"

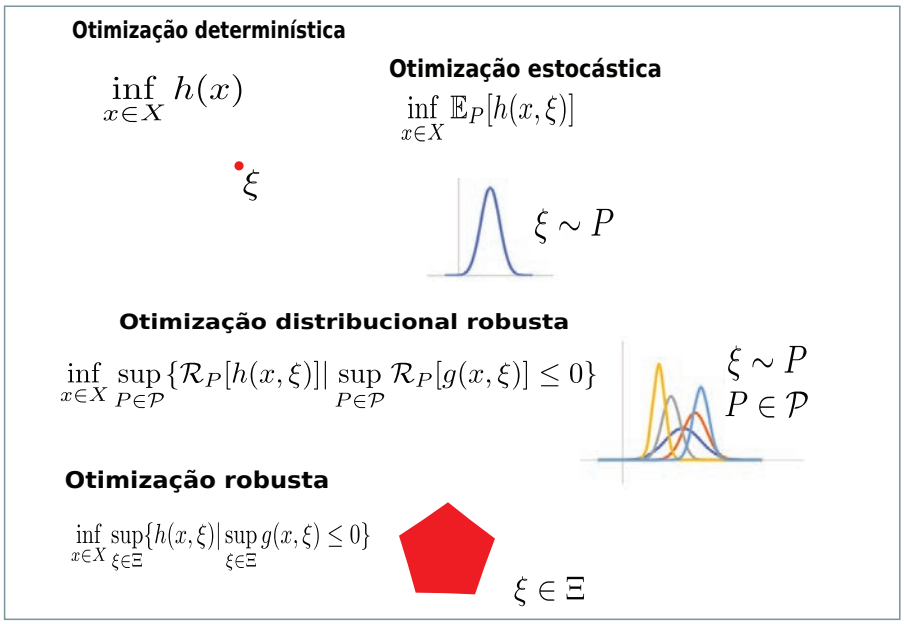


Fig.1 - Comparação entre as formulações de otimização sob incerteza. Adaptado de [5].

cial da distribuição, enquanto (SO) assume conhecimento completo da distribuição.

A otimização distribucional robusta tem recebido recentemente atenção significativa nas comunidades de pesquisa de operações e aprendizado estatístico, e é uma ferramenta atraente para melhorar os modelos de *machine learning*. Isto é porque em vez de escolher um modelo para minimizar o risco empírico, como fazem os algoritmos tradicionais de *machine learning*, estabelece-se um adversário que pode perturbar a distribuição da amostra dentro de um conjunto U centrado em torno do distribuição empírica [9].

As formulações discutidas até aqui também abrangem o contexto temporal de modelos de otimização estocástica sendo de dois estágios, como já discutido, ou multi-estágios, desde que o importante requisito de não antecipação (requisito de que as decisões devem ocorrer antes das observações) seja devidamente tratado [7].

A Figura 1 provê um esquema de comparação entre as formulações de otimização sob incerteza aqui discutidas.

MÉTODOS DE RESOLUÇÃO

Assim como na otimização determinística, não existe um único método de solução que funcione bem para todos os problemas.

Uma direção geral é discretizar a distribuição de probabilidade. Para isso, são utilizadas técnicas de amostragem. Uma das técnicas mais utilizadas para amostragem de uma distribuição de probabilidade é a de Monte Carlo, que é baseada em um pseudo-gerador aleatório usado para aproximar uma distribuição uniforme.

No entanto, métodos Monte Carlo simples podem resultar em grandes erros (intervalos de confiança) e variância. Técnicas de redução de variância existem para reduzir o impacto nas estimativas de Monte Carlo (por exemplo, [2]).

Os métodos tradicionais de resolução desenvolvidos adotam decomposição e múltiplas iterações para resolução do problema e incluem, entre muitos, aproximação média amostral [8], aproximação estocástica [11] e *hedging* progressivo [12].

Atualmente, os métodos de resolução têm sido desenvolvidos com uma natureza altamente interdisciplinar e integrando diferentes técnicas e paradigmas para resolver problemas complexos do mundo real. Os problemas são resolvidos com métodos exatos, algoritmos heurísticos, ou uma combinação deles. Uma abordagem que merece destaque é a de *reinforcement learning* para otimização estocástica. *Reinforcement learning* e otimização estocástica oferecem uma estrutura única que pode modelar qualquer problema de decisão sequencial usando cinco componentes principais: variáveis de estado, variáveis de decisão, variáveis de informação exógenas, função de transição e função objetivo. Quando aplicado a problemas de otimização, o *reinforcement learning* pode ser visto como uma estratégia de busca heurística [10]. Nesse contexto, a abordagem orientada por dados tem ganho destaque devido ao seu fornecimento de insights para tomada de decisão e devido à fácil implementação.

APLICAÇÕES

O crescimento mundial do e-commerce provocou o aumento da importância e valor competitivo da entrega de última milha (*last mile delivery*), e tem levado muitas empresas a buscarem soluções inovadoras para lidar não só com o alto custo envolvido, mas também com os problemas ambientais decorrentes. Como aplicação de otimização sob incerteza, desenvolvemos um estudo em que uma empresa não possui apenas uma frota de veículos disponíveis para fazer entregas, mas também pode utilizar os serviços de motoristas ocasionais (ODs) que estão dispostos a fazer entregas usando seus veículos em troca de uma pequena compensação. Sob esse modelo de negócios, também conhecido como *crowd sourcing*, a empresa procura fazer todas as entregas ao custo total mínimo [13].

Mais recentemente, consideramos um ambiente de entrega de última milha estocástico e dinâmico em que pedidos de entrega, bem como ODs dispostos a fazer entregas, chegam durante todo o dia.

Desenvolvemos uma nova abordagem de aprendizado de reforço profundo (*deep reinforcement learning*, DRL) para o problema que pode lidar com grandes instâncias de problemas da vida real, com uma preocupação na qualidade da solução fornecida. Para oferecer boas soluções, um desafio associado ao desenho da abordagem DRL diz respeito ao espaço de ação, devido `a sua natureza combinatória. Algoritmos de aprendizado por reforço normalmente requerem um espaço de ação que é pequeno o suficiente para enumerar.

Apresentamos um abordagem diferente onde formulamos o problema de seleção de ações como um programa de otimização inteira. Combinamos a estrutura combinatória do espaço de ação com a arquitetura neural da função de valor aprendido, envolvendo técnicas de *machine learning* e otimização inteira (veja [1]).

REFERÊNCIAS

[1] *Anderson, J. Huchette, C. Tjandraatmadja, and J. P. Vielma. Strong mixed-integer programming formulations for trained neural networks, 2018.*

[2] *Arouna. Adaptative monte carlo method, a variance reduction technique. 10(1):1–24, 2004.*

[3] *Bertsimas, D. B. Brown, and C. Caramanis. Theory and applications of robust optimization. SIAM Review, 53(3):464–501, 2011.J. R. Birge and F. Louveaux. Introduction to Stochastic Programming. Springer-Verlag, New York, NY, USA, 1997.*

[4] *Chen and I. C. Paschalidis. Distributionally robust learning. Foun- dations and Trends in Optimization, 4(1-2):1–243, 2020.*

[5] *B. Dantzig. Linear Programming under Uncertainty. Management Science, 1(3-4):197–206, 04-07 1955.*

[6] *Dupa˘cová. Multistage stochastic programs: The state-of-the-art and selected bibliography. Kybernetika, 31(2):151–174, 1995.*

[7] *Kim, R. Pasupathy, and S. Henderson. A Guide to Sample Average Approximation, volume 216, pages 207–243. 01 2015.*

[8] *Lin, X. Fang, and Z. Gao. Distributionally robust optimization: A review on theory and applications. Numerical Algebra, Control and Optimization, 12(1):159–212, 2022.*

[9] *B. Powell. From reinforcement learning to optimal control: A unified framework for sequential decisions. ArXiv, abs/1912.03513, 2019.*

[10] *Robbins and S. Monro. A Stochastic Approximation Method. The Annals of Mathematical Statistics, 22(3):400 – 407, 1951.*

[11] *T. Rockafellar and R. J.-B. Wets. Scenarios and policy aggregation in optimization under uncertainty. Mathematics of Operations Research, 16(1):119–147, 1991.*

[12] *Silva, J. P. Pedroso, A. Viana, and X. Klimentova. A Branch-Price- And-Cut Algorithm for Stochastic Crowd Shipping Last-Mile Delivery with Correlated Marginals. In M. Mu¨ller-Hannemann and F. Perea, editors, 21st Symposium on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2021), volume 96 of Open Access Series in Informatics (OASIcs), pages 12:1–12:20, Dags- tuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum fu¨r Informa- tik.*

PROGRAMAÇÃO DE ROTAS NA ÚLTIMA MILHA
O AMAZON LAST MILE ROUTING RESEARCH CHALLENGE

O Amazon Last Mile Routing Research Challenge é uma competição científica promovida pela Amazon e com o apoio técnico do MIT Center for Transportation & Logistics (MIT CTL). Aberto a estudantes e investigadores universitários de todo o mundo que competem individualmente, ou em equipas de até três membros. Os prémios dos vencedores são $100.000 para o primeiro lugar, $50.000 para o segundo, e $25.000 para o terceiro. A competição decorreu entre abril e junho de 2021.

O principal objetivo do Amazon Last Mile Routing Research Challenge é promover a inovação na intersecção das áreas de investigação operacional e da ciência dos dados. Para o efeito, os participantes são convidados a desenvolver abordagens inovadoras destinadas a resolver problemas reais de roteamento de veículos na última milha.

Apesar dos tremendos avanços na otimização de rotas nas últimas décadas, permanece uma importante lacuna entre o planeamento teórico de rotas e a execução de rotas na vida real. Tradicionalmente, os problemas de roteamento de veículos concentram-se num único objetivo, seja a minimização da distância ou a minimização do custo. Contudo, na distribuição da última milha, a minimização do custo ou da distância nem sempre é suficiente para determinar uma boa rota, porque uma boa rota pode ser uma rota que também permita encontrar um bom local de estacionamento ou que permita evitar áreas de tráfego congestionadas ou locais onde seja difícil realizar manobras com os veículos de transporte.

Os motoristas experientes têm conhecimento tácito sobre a geografia, infraestruturas e clientes a quem entregam as encomendas. Eles conhecem as condições de trânsito das estradas em malha urbana, quando e onde podem facilmente encontrar estacionamento, que entregas podem ser convenientemente servidas em conjunto, e muitas outras condições que são difíceis, se não impossíveis de formalizar num modelo de otimização. Por conseguinte, esta informação tácita não está contida na maioria dos sistemas de planeamento de rotas utilizados nas empresas, fazendo com que os condutores se desviem frequentemente das sequências de rotas originalmente planeadas. Tendo em conta o seu conhecimento tácito, os condutores seguem uma sequência de rota real desviada da rota planeada, que é potencialmente mais conveniente em condições operacionais reais.

A utilização de abordagens baseadas na aprendizagem para compreender e antecipar estes desvios pode levar a rotas mais eficientes e seguras, operações de entrega na última milha mais sustentáveis, condutores mais satisfeitos e uma maior qualidade de serviço. Este é o principal desafio colocado aos participantes, desenvolver abordagens inovadoras usando inteligência artificial, machine learning, deep learning, e outros métodos não convencionais para produzir soluções para o problema de programação de rotas de veículos na última milha que superem os métodos tradicionais da investigação operacional, em termos de qualidade da solução e do custo computacional.

Durante o período da competição os participantes desenvolvem as suas abordagens utilizando dados históricos reais fornecidos pela Amazon de pedidos de entrega, as suas características e as sequências de entregas realizadas pelos motoristas. Cada uma das sequências é qualificada como sendo de alta,

"APESAR DOS TREMENDOS AVANÇOS NA OTIMIZAÇÃO DE ROTAS NAS ÚLTIMAS DÉCADAS, PERMANECE UMA IMPORTANTE LACUNA ENTRE O PLANEAMENTO TEÓRICO DE ROTAS E A EXECUÇÃO DE ROTAS NA VIDA REAL"

ANTÓNIO GALRÃO RAMOS

INESC TEC e Instituto Superior de Engenharia do Porto
agr@isep.ipp.pt

MANUEL PEREIRA LOPES

INESC TEC e Instituto Superior de Engenharia do Porto
mpl@isep.ipp.pt

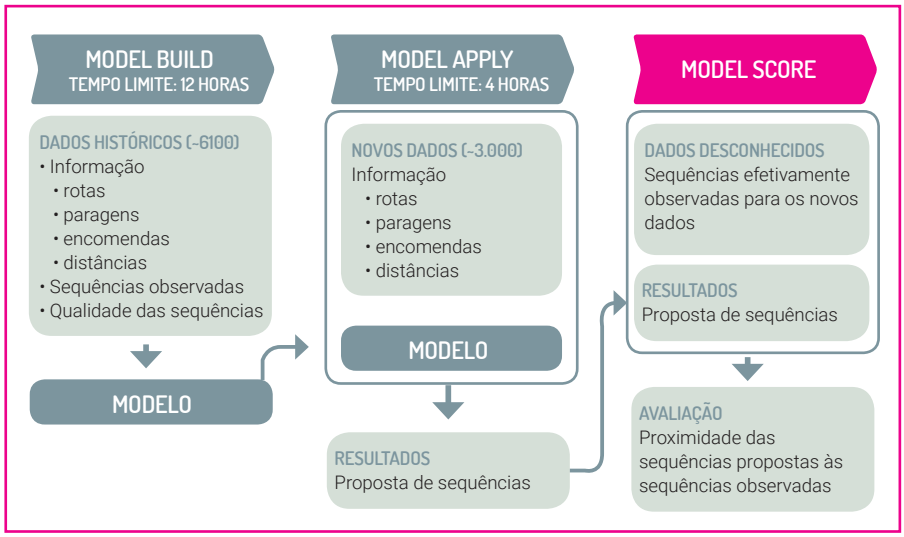


Fig 1- Fases do processo

"A UTILIZAÇÃO DE ABORDAGENS BASEADAS NA APRENDIZAGEM PARA COMPREENDER E ANTECIPAR ESTES DESVIOS PODE LEVAR A ROTAS MAIS EFICIENTES E SEGURAS, OPERAÇÕES DE ENTREGA NA ÚLTIMA MILHA MAIS SUSTENTÁVEIS, CONDUTORES MAIS SATISFEITOS E UMA MAIOR QUALIDADE DE SERVIÇO"

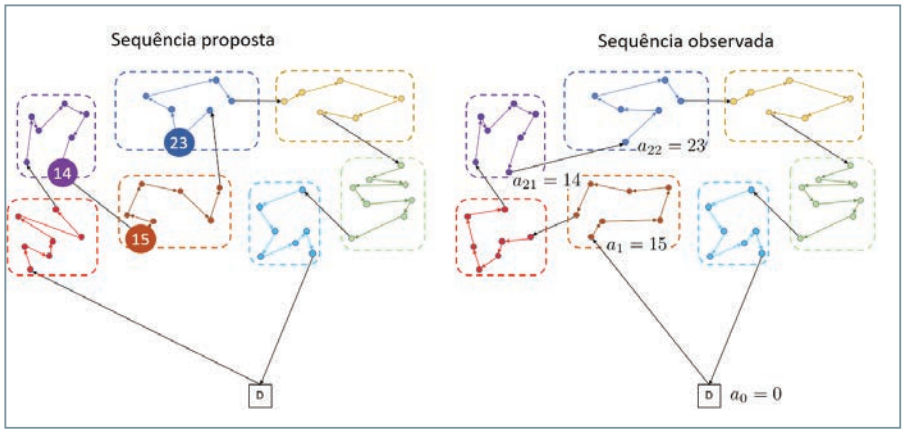


Fig 2 - Desvio da sequência proposta

média ou baixa qualidade. No final do período de competição as abordagens propostas são avaliadas usando um novo conjunto de dados, para os quais devem ser geradas sequências de rotas que funcionem bem num determinado conjunto de métricas de desempenho.

O processo tem três fases (Figura 1):

1- Model Build: construção de um modelo com base em dados históricos.

2- Model Apply: propor rotas de alta qualidade para novos dados.

3- Model Score: avaliação com base nos desvios das rotas propostas relativamente às rotas de alta qualidade observadas na realidade.

Com o apoio técnico de cientistas do MIT CTL, a Amazon avalia as submissões com base numa combinação das seguintes três métricas: precisão da solução (isto é, a proximidade das sequências das rotas planeadas para as sequências das rotas reais de alta qualidade), tempo total da rota (como medida de eficiência da rota) e tempo total de computação.

A métrica de precisão das soluções geradas é realizada em primeiro lugar rota a rota e posteriormente agregada numa pontuação de submissão. A componente principal na determinação da qualidade da rota é a componente de Desvio da Sequência Observada (DS) (1), em que n representa o número total de paragens numa sequência e a_i a posição na sequência proposta da posição i na sequência observada. Esta métrica pode ser vista como uma variação percentual entre a sequência proposta e a observada, sendo que um DS igual a zero representa um ajuste perfeito entre ambas. A Figura 2 ilustra a determinação dos valores de a_i. Por exemplo, na rota real a primeira paragem corresponde à 15ª paragem na rota proposta o que se traduz num a_1 =15.

$$DS = \frac{2}{n(n-1)} \sum_{i=1}^{n} |a_i - a_{i-1}| - 1 \qquad (1)$$

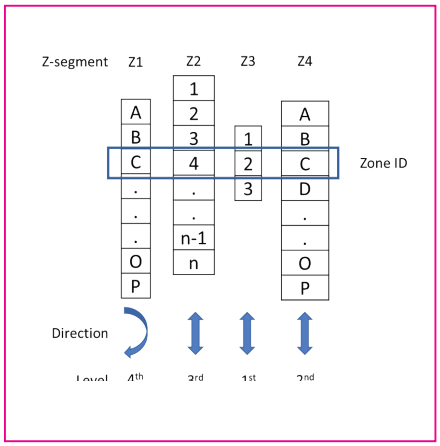


Fig 3 - Gerador de sequências de Zone IDs

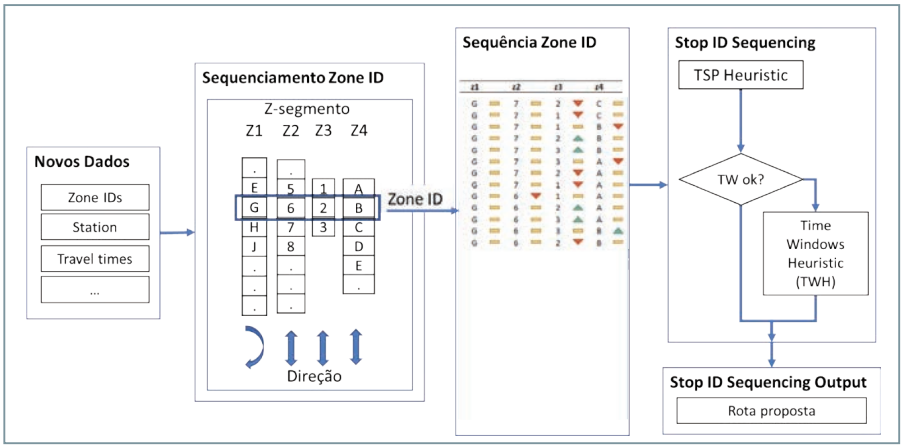


Fig 4 - Estrutura global do modelo

ABORDAGEM PROPOSTA

A estratégia geral definida para abordar o desafio, foi naturalmente condicionada pelas regras da competição. O desafio consiste em desenvolver uma solução a ser executada em três etapas: na fase de Model Build, o modelo é treinado utilizando dados históricos (6112 rotas), na fase de Model Apply as sequências de rotas são determinadas para novos dados (3052 rotas) e na fase Model Score a qualidade do ajustamento das sequências de rotas geradas na fase de Model Apply é medido em relação às sequências de rotas efetivamente realizadas.

A partir da análise preliminar dos dados, verificou-se que a abordagem seguida no planeamento de rotas utilizado pela Amazon está de acordo com as abordagens aos problemas de roteamento de veículos em problemas de distribuição em larga escala, tais como os da última milha, em que primeiro a nível tático o território é dividido em zonas de distribuição padrão e depois a nível operacional o roteamento é realizado em duas fases, na primeira fase as parcelas são agrupadas por veículo em função da procura, e numa segunda fase é realizada a sequenciação das paragens da rota para cada veículo.

A equipa MEGI do ISEP, propôs uma solução baseada num algoritmo de sequenciamento das zonas de distribuição padrão (Zone ID) identificadas na fase exploratória dos dados. Cada Zone ID é identificada por um conjunto de carateres alfanuméricos e não alfanuméricos. Com base nos dados históricos o campo Zone ID foi segmentado utilizando os caracteres não alfanuméricos como delimitadores e o tipo de carácter alfanumérico. Isso resultou em quatro novos campos que designámos como Z1, Z2, Z3 e Z4 e na determinação das suas relações hierárquicas. Estes novos campos são doravante referidos como segmentos Z.

A ideia geral do algoritmo é a seguinte: a sequência de Zone IDs é gerada movendo uma unidade de um segmento Z de cada vez. Cada vez que um segmento Z realiza uma deslocação completa numa determinada direção, o próximo segmento Z de nível superior move uma unidade numa determinada direção. Em seguida, volta ao segmento Z de nível inferior e, movendo uma unidade de cada vez, realiza uma deslocação completa na direção oposta, num movimento pendular (Figura 3).O algoritmo de Zone ID requer como entradas a hierarquia dos segmentos Z, a direção de deslocação para cada um dos segmentos Z e os identificadores da zona da rota. A hierarquia dos segmentos Z e a direção de deslocação para os segmentos Z de uma dada rota são obtidas durante a fase de aprendizagem do modelo (Model Build). O resultado do algoritmo de sequenciamento de Zone IDs é usado por uma heurística de sequenciamento de entregas (Heurística TSP) para gerar as rotas propostas (Figura 4).

"NA COMPETIÇÃO PARTICIPARAM 229 EQUIPAS, NUM TOTAL DE 2285 PARTICIPANTES, DAS QUAIS PASSARAM À FASE FINAL 45 EQUIPAS E 107 PARTICIPANTES, DE 71 INSTITUIÇÕES E 29 NACIONALIDADES (EUA, CANADÁ, ALEMANHA, FRANÇA, UK, BÉLGICA, PAÍSES BAIXOS,.... A EQUIPA MEGI DO ISEP ALCANÇOU O 4º LUGAR"

RESULTADOS

As experiências computacionais foram executadas num CPU Intel Xeon Gold 6148 2,40GHz com 96 GB de RAM. Todos os algoritmos foram implementados em Python 3.9. A Figura 5 apresenta a frequência das pontuações obtidas para as rotas das instâncias. Pode-se observar que 90% dos resultados são inferiores a 0,09, e 18,2% das rotas propostas são quase uma correspondência perfeita com as sequências das rotas reais correspondentes.

Na competição participaram 229 equipas, num total de 2285 participantes, das quais passaram à fase final 45 equipas e 107 participantes, de 71 instituições e 29 nacionalidades (EUA, Canadá, Alemanha, França, UK, Bélgica, Países Baixos, Dinamarca, Noruega, Austrália, China, India, Indonésia, Coreia do Sul, Brasil, Peru, Egito, entre outros).

A equipa MEGI do ISEP alcançou o 4º lugar da competição (Figura 6), que foi ganha por uma equipa de três universidades (University of Waterloo, Canada, University of Bonn, Alemanha, e Roskilde University, Dinamarca), 2º lugar para uma equipa do MIT, e o 3º lugar para uma equipa das universidades HEC Montréal, Canada, e New South Wales, Austrália.

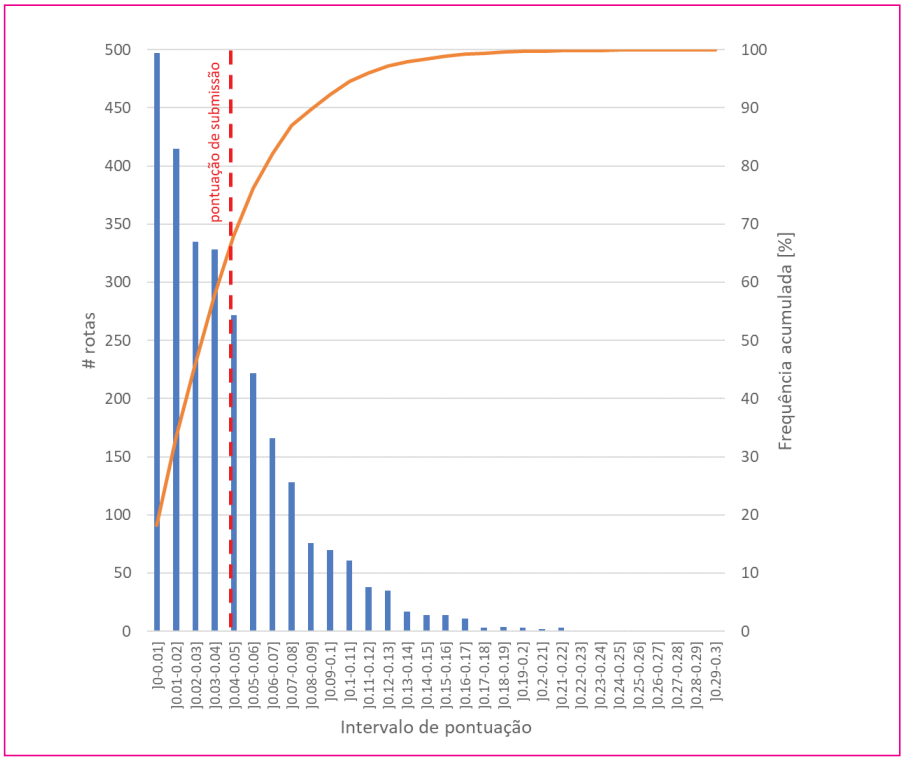


Fig 5 - Pontuação das rotas

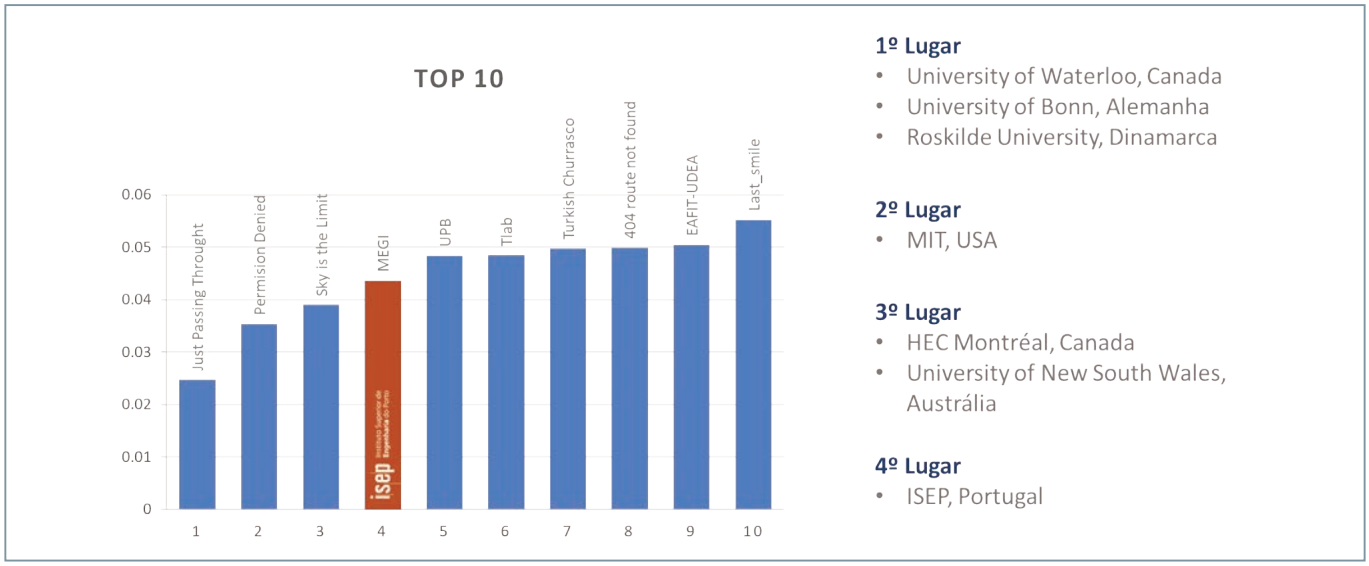


Fig 6 - Resultados finais

(DIS)SIMILARIDADE EM PROCESSOS DE TOMADA DE DECISÃO COM MÚLTIPLOS CRITÉRIOS E CATEGORIAS NOMINAIS

ANA SARA COSTA

CEGIST, Instituto Superior Técnico, Universidade de Lisboa
anasaracosta@tecnico.ulisboa.pt

ENQUADRAMENTO

A diversidade de problemas de classificação em processos de tomada de decisão tem incentivado os investigadores a desenvolver métodos e técnicas de apoio à decisão direcionados para a criação de modelos de decisão para apoiar e permitir lidar melhor com tais situações. Em geral, este tipo de problemas envolve múltiplos critérios para avaliar os objetos de decisão e a necessidade de afetá-los a categorias predefinidas, de acordo com as preferências do decisor. Quando as categorias têm uma ordem de preferência entre elas, estamos na presença de problemas de classificação ordinal, caso contrário, consideramos tratar-se de um problema de classificação nominal. Tendo presente o estado da arte neste domínio, exploramos alguns aspetos, com vista a tornar os métodos mais adequados para problemas multicritério de classificação nominal [1], nomeadamente a consideração de julgamentos subjetivos de similaridade e dissimilaridade, e de efeitos de interação entre critérios. Este é o ponto de partida para o desenvolvimento de novos métodos, adotando uma abordagem integrada, e com aplicabilidade em diversos contextos atuais.

MÉTODO PROPOSTO

Os métodos multicritério de classificação nominal pressupõem que os objetos de decisão (aqui designados por ações, de acordo com Roy [2]) são avaliados considerando diferentes pontos de vista para serem afetados a categorias definidas a priori e sem qualquer ordem de preferência entre elas (aqui designadas por categorias nominais).
O método CAT-SD (*CATegorization by Similarity-Dissimilarity*) [3] foi desenvolvido com base em conceitos de similaridade e dissimilaridade entre duas ações, particularmente no modelo de contraste de Tversky [4]. O método está representado, de forma simplificada na Fig. 1, incluindo-se os dados e parâmetros (*input*), a formulação matemática e os resultados do procedimento de afetação (*output*).

Os dados do problema de classificação são as ações, os critérios e as categorias nominais (no método CATSD considera-se uma categoria adicional *dummy* para aceitar ações que não sejam consideradas adequadas para as restantes categorias). Conhecendo os desempenhos de cada ação em cada critério, constrói-se uma tabela de desempenhos. Para cada categoria, os parâmetros de preferência definidos pelo decisor incluem os pesos dos critérios (representam a importância relativa dos mesmos), os coeficientes de interação entre dois critérios (efeitos de interação de mútuo fortalecimento, mútuo enfraquecimento e antagonismo), caso existam, as ações de referência (ações representativas que caracterizam as categorias, sendo necessário definir pelo menos uma por categoria) e os limiares de semelhança. Para cada critério é construída uma função de similaridade-dissimilaridade (função SD) que permite modelar a similaridade e dissimilaridade entre uma ação a ser avaliada e uma ação de referência, ao longo da escala do critério. Considera-se a diferença entre os níveis de desempenho das duas ações e, de acordo com os julgamentos do decisor, é avaliada a presença de similaridade ou dissimilaridade entre elas. A função SD é positiva quando se considera que existe similaridade, e é negativa quando se considera que existe dissimilaridade, variando entre -1 e 1.
No método CAT-SD, cada ação a categorizar

"ESTE MÉTODO CONSIDERA OS JULGAMENTOS DO DECISOR NO QUE DIZ RESPEITO À SIMILARIDADE E DISSIMILARIDADE ENTRE AÇÕES,PERMITE MODELAR INTERAÇÕES ENTRE DOIS CRITÉRIOS (...)"

"O DECSPACE TEM SIDO DESENVOLVIDO PARA OFERECER UMA GRANDE VARIEDADE DE MÉTODOS, NUMA INTERFACE DE FÁCIL UTILIZAÇÃO (...)"

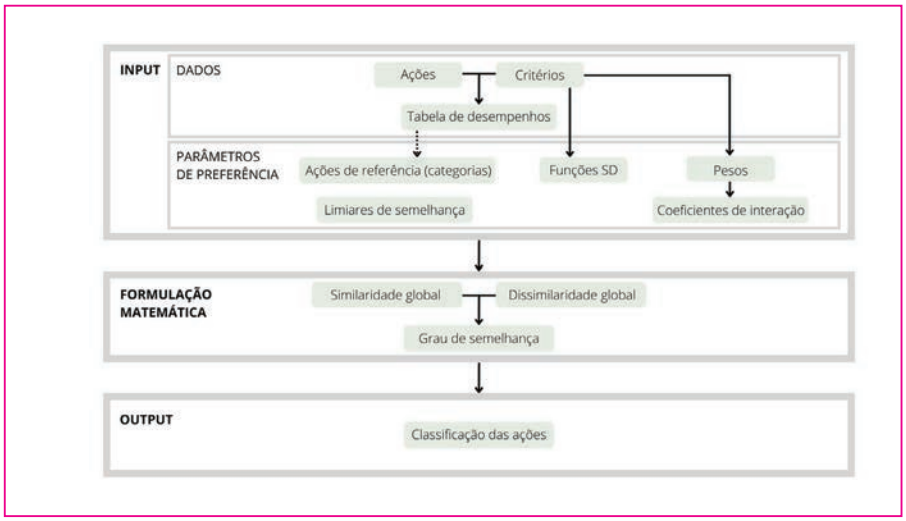


Fig. 1 - Diagrama do método CAT-SD (adaptado de [5]).

é comparada com as ações de referência e, considerando todos os parâmetros de preferência referidos, é possível obter-se um grau de semelhança global entre cada ação e cada categoria, que varia entre 0 e 1. Este grau de semelhança global é comparado com o limiar de semelhança definido para cada categoria e cada ação é afetada a pelo menos uma categoria. De acordo com o procedimento de afetação do método, uma ação é atribuída a uma dada categoria se existir um grau de semelhança entre a ação e o conjunto de ações de referência dessa categoria que seja pelo menos igual ao limiar de semelhança da categoria e, caso isso não se verifique para nenhuma categoria, a ação é afetada à categoria *dummy*.

Assim, este método considera os julgamentos do decisor no que diz respeito à similaridade e dissimilaridade entre ações, permite modelar interações entre dois critérios e definir diferentes conjuntos de parâmetros entre as categorias.

O método tem potencial de aplicação em diversas situações que envolvam a classificação em categorias nominais com múltiplos critérios, em várias áreas e contextos e, como base nas recomendações elaboradas para os decisores, poderá trazer valor para as decisões nas suas organizações.

IMPLEMENTAÇÃO NO DECSPACE

O método CAT-SD está implementado num sistema *web*, designado DECSPACE[1], que disponibiliza métodos multicritério de apoio à decisão a qualquer utilizador. O DECSPACE tem sido desenvolvido para oferecer uma grande variedade de métodos, numa interface de fácil utilização, com possibilidade de *customization*, potenciando a sua utilização por decisores de diversas áreas, sem que lhes sejam exigidas competências técnicas sobre como utilizar os métodos. Assim, a utilização do DECSPACE pode ter quer fins de investigação e ensino quer um propósito profissional (por exemplo, projetos de consultoria).

O DECSPACE apresenta uma interface intuitiva que permite criar projetos com um ou mais destes métodos em simultâneo, através da construção de um *workflow*, combinando os principais componentes dos métodos através de "caixas" devidamente ligadas entre si (de forma semelhante ao *diviz*[2]). A Fig. 2 apresenta o *workflow* de um modelo CAT-SD construído no DECSPACE, encontrando-se do lado esquerdo da "caixa CAT-SD" todos os elementos de *input* do método, designada-

1 - http://app.decspacedev.sysresearch.org

2 - https://www.diviz.org

mente os dados e parâmetros do método, e do lado direito o *output* que contém os resultados do modelo construído, os quais podem ser exportados. Todos os dados podem ser importados (por exemplo, em ficheiros CSV) ou inseridos manualmente nas tabelas devidamente identificadas dentro da "caixa CAT-SD", conforme ilustra a Fig. 3.

APLICAÇÃO NO EXÉRCITO PORTUGUÊS

Em colaboração com o Exército Português através do Centro de Psicologia Aplicada do Exército, com o objetivo de avaliar a adequabilidade dos candidatos militares para as várias especialidades das Tropas Especiais do Exército, tendo em conta as características dos perfis exigidos, foi aplicado o método CAT-SD [6]. O decisor é um militar especialista na avaliação e seleção de candidatos para o Exército. Neste estudo, assume-se que os candidatos completaram a formação militar inicial e pretendem seguir uma das seguintes especialidades: Comandos, Paraquedistas, Operações Especiais (OE) e Snipers. Os snipers estão incluídos nas OE, mas são forças de elite com um perfil muito específico, pelo que se considerou a categoria "snipers" separadamente. Assim, para além da categoria *dummy* "não atribuídos" definiram-se mais quatro categorias distintas, conforme se observa na Fig. 4, a qual ilustra a perspetiva geral da situação de decisão em análise.

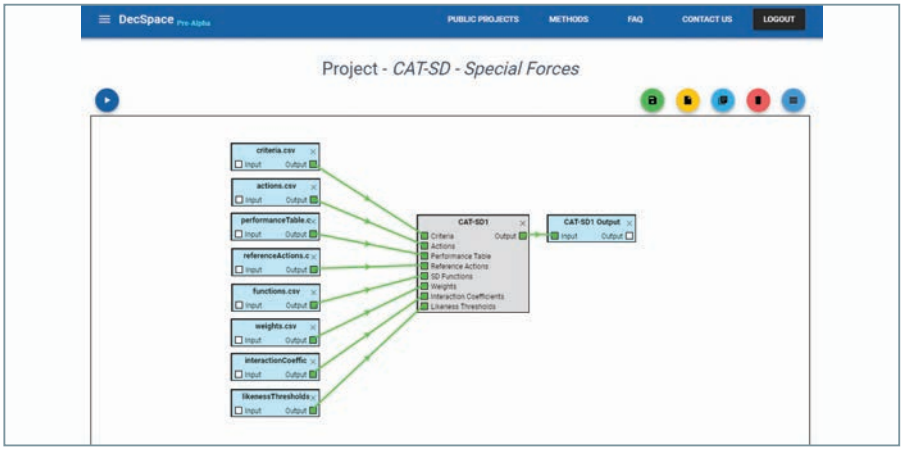
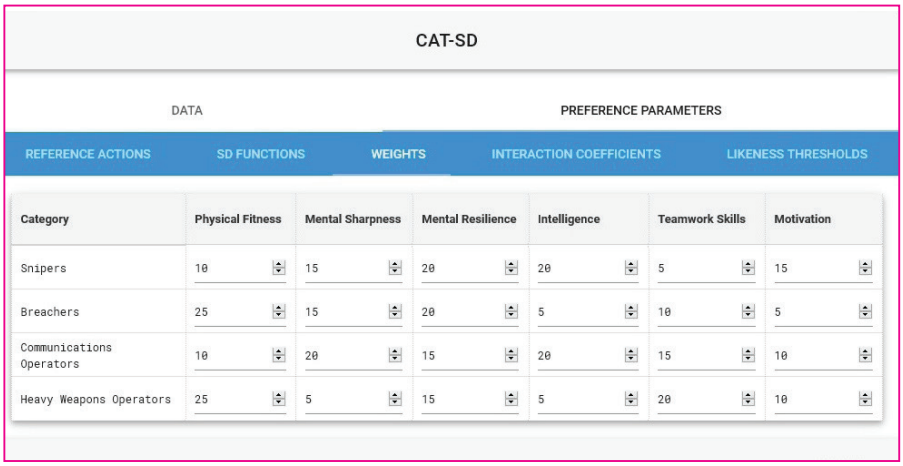


Fig. 2 - *Workflow* do método CAT-SD no DECSPACE [6].

"DE ACORDO COM O DECISOR, OS RESULTADOS SÃO SATISFATÓRIOS E COERENTES COM OS REQUISITOS PARA O RECRUTAMENTO E SELEÇÃO DOS MILITARES DAS TROPAS ESPECIAIS DO EXÉRCITO"



CAT-SD

DATA | PREFERENCE PARAMETERS

REFERENCE ACTIONS | SD FUNCTIONS | WEIGHTS | INTERACTION COEFFICIENTS | LIKENESS THRESHOLDS

Category	Physical Fitness	Mental Sharpness	Mental Resilience	Intelligence	Teamwork Skills	Motivation
Snipers	10	15	20	20	5	15
Breachers	25	15	20	5	10	5
Communications Operators	10	20	15	20	15	10
Heavy Weapons Operators	25	5	15	5	20	10

Fig. 3 - Exemplo das tabelas de dados do método CAT-SD no DECSPACE [5].

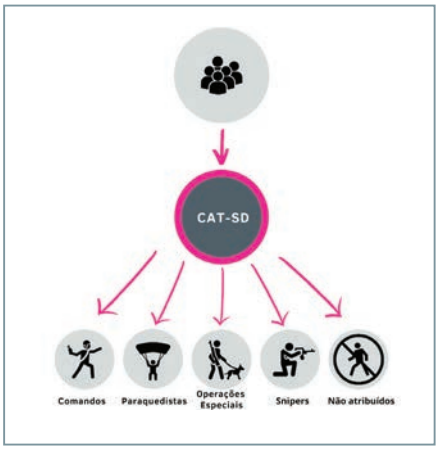


Fig. 4 - Visão geral do caso de categorização das Tropas Especiais do Exército.

Fig. 5 - Critérios utilizados para avaliação dos candidatos militares no estudo de Costa et al. [6].

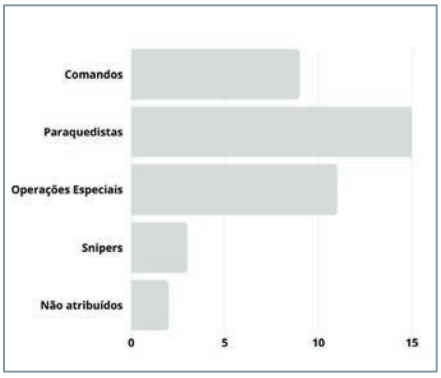


Fig. 6 - Número de candidatos atribuídos por categoria.

Em geral, o perfil representativo de um sniper apresenta valores de desempenho nos critérios mais elevados, bem como um limiar de semelhança mais alto (neste estudo, os dados são fictícios por razões de confidencialidade). Isto significa que é exigido um maior grau de semelhança entre um potencial candidato e este perfil de referência para ser afetado à categoria, ou seja, ser considerado adequado para a especialidade.

Considerando uma abordagem construtiva, através da interação do analista e do decisor, no final, foram selecionados nove critérios de avaliação: robustez física, inteligência, aptidão numérica, aptidão espacial, aptidão para a mecânica, velocidade percetiva, aptidão psicomotora, personalidade e condição médica (Fig. 5). Estes critérios consideram os resultados das provas de avaliação física, psicológica e médica, sendo que alguns deles usam escalas percentílicas e outros utilizam julgamentos qualitativos associados a diferentes níveis de desempenho em escala ordinal.

Os parâmetros do modelo foram obtidos com base no diálogo entre o analista e o decisor, seguindo protocolos de interação previamente elaborados, no sentido de facilitar o processo de determinação dos parâmetros. Com o intuito de se determinar o valor dos pesos dos critérios recorreu-se ao método das cartas SRF [7], obtendo-se diferentes conjuntos de pesos entre as várias especialidades.

Os resultados da afetação foram obtidos através do DECSPACE. Assim, encontramos para cada candidato uma categoria (ou categorias) adequada(s) de especialidade nas Tropas Especiais do Exército. Alguns candidatos foram atribuídos a uma categoria específica e outros foram considerados adequados para mais do que uma categoria. Numa visão geral, como mostra a Fig. 6, verifica-se que a maioria dos candidatos foi considerada adequada para Paraquedista (15 de 20), nove para Comandos, onze para OE, e apenas três dos candidatos tinha um perfil compatível com o exigido para Sniper (estes foram considerados igualmente aptos para qualquer uma das outras especialidades, como era expectável, segundo o decisor). Somente dois candidatos não foram considerados aptos para nenhuma das especialidades.

De acordo com o decisor, os resultados são satisfatórios e coerentes com os requisitos para o recrutamento e seleção dos militares das Tropas Especiais do Exército. Este estudo apresenta-se, assim, como uma contribuição relevante para a investigação dos processos de recrutamento e seleção do Exército Português.

"O MÉTODO TEM POTENCIAL DE APLICAÇÃO EM DIVERSAS SITUAÇÕES QUE ENVOLVAM A CLASSIFICAÇÃO EM CATEGORIAS NOMINAIS COM MÚLTIPLOS CRITÉRIOS, EM VÁRIAS ÁREAS E CONTEXTOS (...)"

REFERÊNCIAS

[1] *Zopounidis C., & Doumpos M. (2002). Multicriteria classification and sorting methods: A literature review. European Journal of Operational Research, 138(2), 229–246.*

[2] *Roy B. (1996) Multicriteria Methodology for Decision Aiding. Kluwer Academic Publishers.*

[3] *Costa, A. S., Figueira, J. R., & Borbinha, J. (2018). A multiple criteria nominal classification method based on the concepts of similarity and dissimilarity. European Journal of Operational Research, 271(1), 193–209.*

[4] *Tversky, A. (1977). Features of similarity. Psychological Review, 84(4), 327–352.*

[5] *Costa, A. S. (2020) A multiple criteria integrated approach for nominal classification problems: Methods and applications [Tese de Doutoramento, Instituto Superior Técnico – Universidade de Lisboa]. Sotis. https://sotis.tecnico.ulisboa.pt/record/cf3f8ee8-7517-4b3e-b7c5-15ce188f47fd*

[6] *Costa, A. S., Figueira, J. R., & Borbinha, J. (2021). A multiple criteria socio-technical approach for the Portuguese Army Special Forces recruitment. 4OR - A Quarterly Journal of Operations Research. https://doi.org/10.1007/s10288-021-00481-2*

[7] *Figueira, J. R., & Roy, B. (2002). Determining the weights of criteria in the ELECTRE type methods with a revised Simos' procedure. European Journal of Operational Research, 139(2), 317–326.*

A IMPORTÂNCIA DA INVESTIGAÇÃO OPERACIONAL NUMA EMPRESA DE LOGÍSTICA PARA E-COMMERCE EM RÁPIDA EXPANSÃO PELA EUROPA

BRUNO VIEIRA

Analista de Investigação Operacional
na Paack
brunovieira3@hotmail.com

Começo por agradecer o convite da APDIO para partilhar a minha experiência, maioritariamente internacional, no desenvolvimento e aplicação de técnicas de IO em distintas áreas, e mais recentemente no mercado de logística para comércio eletrónico onde tenho estado bastante ativo desde o último meio ano.

Nascido e criado no Porto, comecei a minha experiência profissional como investigador de IO no INESC TEC, em 2012, enquanto estudava o Mestrado de Engenharia Eletrotécnica - Sistemas e Planeamento Industrial no ISEP. Pela mão da excelente professora e agora amiga Ana Viana, desde cedo senti uma forte paixão pela área de IO, sobretudo pelo seu potencial para otimizar processos e criar um impacto positivo no mundo ao poder ser aplicado a praticamente todas as áreas que afetam diretamente as pessoas e o planeta. No INESC TEC, participei essencialmente em 2 projetos de otimização. No primeiro, desenvolvi métodos matemáticos com teoria da utilidade para facilitar a integração de energias renováveis no problema de geração de energia elétrica considerando a variabilidade inerente à energia eólica. No segundo projeto, construí um modelo híbrido que combinava uma metaheurística (BRKGA) e um solver (CPLEX) para resolver problemas de sequenciamento de operações como os encontrados, por exemplo, no movimento de contentores em terminais portuários.

Com um espírito explorador e internacional aliado a um forte desejo de desenvolver e aplicar técnicas de IO na área da saúde, em 2015 comecei um doutoramento no *Netherlands Cancer Institute* (NKI), o maior hospital oncológico dos Países Baixos localizado em Amesterdão, em colaboração com a Universiteit *Twente*. Uma experiência que considero bastante aplicada, já que trabalhava diariamente no departamento de radioterapia do hospital colaborando diretamente com médicos radio-oncologistas e radioterapeutas. Durante os 5 anos de doutoramento, desenvolvi várias ferramentas baseadas em IO para a otimização logística de processos de planificação em radioterapia, sempre com o objetivo de aumentar a qualidade dos tratamentos e o bem-estar dos pacientes. Entre esses métodos encontram-se um modelo matemático para a alocação ótima de radioterapeutas às diversas operações que executam, um modelo de simulação para quantificar as possíveis consequências de diferentes políticas de marcação da data de início de tratamento, a um programa também baseado em otimização matemática para melhorar o processo de agendamento das várias sessões diárias de radioterapia considerando janelas temporais indicadas pelos próprios pacientes. Em 2020, decidi mudar-me para Barcelona para, em colaboração com a *Universitat Pompeu Fabra* como investigador pós-doutoral, desenhar e implementar (através de um website dedicado) uma heurística construtiva para otimizar a planificação combinada de serviços sociais e de saúde ao domicílio, de extrema importância para a crescente população mais idosa. É gratificante verificar que o sistema está atualmente em fase de testes com vista à sua utilização regular por parte dos respetivos decisores.

Em meados de 2021, ao fim de 9 anos de investigação, decidi começar uma nova aventura no mundo industrial ao juntar-me à Paack, uma *scale-up* na área de logística para *e-commerce*. Com uma visão de liderar a entrega sustentável de encomendas na Europa, Paack apresenta uma atitude positiva no aspeto ambiental ao assegurar uma percentagem elevada de entregas em veículos elétricos, aliada a um serviço de qualidade concentrado na experiência positiva do consumidor final, que tem, por exemplo, a possibilidade de escolher a hora de entrega da sua encomenda no mesmo dia ou no dia seguinte à compra *online*.

Já presente em Portugal, Espanha, Reino Unido, França e Itália, a empresa de apenas 6 anos sediada em Barcelona está em rápida expansão dentro destes países e em busca de novos mercados na Europa. Nesta fase, a necessidade de aplicar conhecimento de IO é de extrema importância para o desenho de um plano de expansão adequado, eficiente e sustentável. Os aspetos estratégicos são cruciais e, como tal, o grupo de investigação operacional ao qual pertenço tem naturalmente estado focado em problemas estratégicos de desenho de rede de distribuição, incluindo a localização e dimensionamento de centros de distribuição e a definição de linhas de transporte de primeira e média milha, normalmente efetuado por camiões. Com um horizonte temporal de médio/longo prazo e sempre balanceando custos com qualidade de serviço, atualmente utilizamos otimização matemática e simulação para dar suporte analítico a decisões de caráter estratégico tanto nos países atuais como em potenciais novos mercados.

A nível operacional, é também fascinante identificar a enorme quantidade de problemas de otimização combinatória nos quais a IO, e em particular a otimização matemática, podem ser extremamente úteis. Desde a alocação semanal de volume (encomendas) a diferentes frotas de última milha, algoritmos de roteamento diário dos próprios veículos de entrega final, ao dimensionamento da mão-de-obra diária necessária nos armazéns para processar encomendas, os problemas de decisão de elevada complexidade estão presentes em quase todas as etapas do processo de entrega, com a dificuldade acrescida associada à variabilidade do número de encomendas a processar diariamente e à existência de janelas temporais para entrega.

Tem sido bastante satisfatório experienciar o elevado valor trazido pelas técnicas de IO na resolução de problemas de decisão reais e atuais, e da minha parte continuarei a fazer o possível por demonstrar o valor que estas podem trazer às organizações e com elas criar um impacto positivo no mundo e na sociedade.

EVENTOS REALIZADOS

2º Workshop EUROYoung

O segundo workshop EUROYoung foi realizado em Portugal (Porto), nas instalações da FEUP e do INESC TEC, tendo contado com a parceria da EURO, FEUP, INESC TEC, APDIO e LTPlabs. Este evento juntou jovens investigadores de 10 países europeus, num ambiente de partilha e networking, com um copo de vinho do Porto a acompanhar a atividade social.
O EUROYoung é um Fórum da EURO (The Association of European Operational Research Societies) dedicado aos investigadores que são novos no mundo da Investigação Operacional (IO). O Fórum nasce com os seguintes objetivos:

- Fomentar a colaboração entre estudantes e investigadores do IO em início de carreira;
- Proporcionar aos investigadores em início de carreira ferramentas que lhes permitam avançar nas suas carreiras, com especial ênfase na formação.
- Criar redes tanto entre os investigadores em início de carreira como com os líderes seniores na área de IO.
- Ligar a procura e a oferta no mercado de trabalho de IO, tanto no meio académico como na indústria.

Mais informações em: https://euroyoung.eu/

EVENTOS A REALIZAR

IO2022 - XXII Congresso da Associação Portuguesa de Investigação Operacional

XXII CONGRESSO DA APDIO
IO 2022
Universidade de Évora . 6-8 Novembro

O IO2022 - XXII Congresso da Associação Portuguesa de Investigação Operacional realiza-se na Universidade de Évora de 6 a 8 de novembro de 2022, sendo uma organização conjunta do Centro de Estudos e Formação Avançada em Gestão e Economia (CEFAGE) e do Centro de Investigação em Matemática e Aplicações (CIMA), em colaboração com a direção da APDIO.

O tema do congresso é «IO em Tempos de Turbulência: Adaptação e Resiliência». Depois de 2 anos a viver em pandemia, vivemos agora tempos de guerra que, para além da crise humanitária, evidenciou problemas graves como a dependência energética e alimentar. Em simultâneo, precisamos de continuar a encontrar soluções para fazer face às alterações climáticas e construir um futuro mais sustentável. São muitos os problemas/desafios a resolver em tempos de turbulência, mas a Investigação Operacional pode ajudar-nos a propor soluções para esses problemas e ajudar a construir organizações/sociedades com capacidade de adaptação e resiliência a situações de crise. O desafio no IO2022 é a apresentação de trabalhos que reflitam o papel da IO na resolução de problemas de organizações e da sociedade, especialmente os defrontados em tempos de turbulência e de crise e onde as soluções incorporem as características de adaptação e resiliência.
Já estão confirmados os seguintes oradores plenários: Sriram Narayanan (Michigan State University); Miguel Araújo (Prémio Pessoa 2018); José Rui Figueira e Filipe Froes (IST, Universidade de Lisboa e Hospital Pulido Valente) e João Silva (Vale da Rosa).
Mais informações em: http://apdio.pt/web/io2022

APDIO
ASSOCIAÇÃO PORTUGUESA DE INVESTIGAÇÃO OPERACIONAL

Associação Portuguesa de Investigação Operacional
Departamento de Engenharia e Gestão
Instituto Superior Técnico
Av. Rovisco Pais, 1
1049-001 Lisboa, Portugal
apdio@euro-online.org

http://apdio.pt/home
https://www.facebook.com/APDIO.PT/
https://www.linkedin.com/in/apdio-pt-545718177/

BOLETIM
APDIO

Equipa Editorial
Elsa Silva
elsa.m.silva@inesctec.pt
Maria João Santos
mjsantos@inesctec.pt
Sara Martins Correia
ssbm@estg.ipp.pt

Design
Inês Assis
inesassis.design@gmail.com
Impressão
Gráfica Pacense, Lda.
Tiragem
200 exemplares