

Contribuições de aprendizado por reforço em escolha de rota e controle semafórico

ANA L. C. BAZZAN¹

Introdução

A CRESCENTE demanda por mobilidade é uma das características da nossa sociedade. Por exemplo, de acordo com a empresa Inrix, o custo dos congestionamentos nos Estados Unidos em 2017 foi de 305 bilhões de dólares, um aumento de 10 bilhões em relação a 2016. No Brasil, de 2004 a 2014, o número de veículos para 1.000 habitantes aumentou de 119,7 para 206,0. Sabe-se que os investimentos na infraestrutura da rede de transporte não acompanham o crescimento da frota veicular; portanto, os congestionamentos são um fenômeno sempre presente que representa grande desafio para a mobilidade urbana. O impacto direto e indireto dos congestionamentos em áreas urbanas e interurbanas é imenso e resulta em custos que podem atingir até 1% do PNB.

Segundo especialistas, tais custos são de vários tipos: o tempo ocioso das pessoas no trânsito (conhecido em economia como custo de oportunidade), e os gastos pecuniários impostos à sociedade, como os gastos referentes ao consumo de gasolina e diesel, bem como o impacto dos poluentes na saúde da população e o aumento no custo do transporte de carga. Por fim, o impacto negativo também se faz sentir na estrutura econômica do país, na saúde, na qualidade de vida e no bolso do cidadão. “Soluções” como pedágios urbanos, rodízio de placas etc., praticadas atualmente no Brasil, são impopulares. O cidadão necessita ver o retorno do seu sacrifício, seja ele monetário ou não. Dessa forma, existe uma grande demanda por soluções que envolvam inteligência e informação como forma de oferecer uma contrapartida à população.

Do ponto de vista prático, a questão sobre como se deslocar de A até B de maneira eficiente é um tema que está na agenda da maioria dos habitantes das cidades do planeta, haja vista o número de aplicativos para auxiliar na escolha de uma rota ou de um transporte público. Uma maneira popular de mitigar esse problema é melhor usar a infraestrutura existente. Felizmente, os avanços científicos e tecnológicos nos permitem, hoje, sermos otimistas em relação a essa tarefa. Pelo lado científico, recentes avanços na pesquisa em Inteligência Artificial (IA) e Aprendizado de Máquina (AM) as colocam como poderosas ferramentas, na medida em que permitem ou alargam os horizontes no que tange à otimiz-

ção do uso dos recursos existentes. Isso vai desde um controle semafórico mais inteligente, até serviços que não apenas indicam rotas menos congestionadas para os usuários do sistema de transporte, mas o fazem de modo a aliviar o uso de certas vias arteriais, balanceando o uso do recurso viário como um todo.

É nesse contexto que surge a agenda em torno de cidades inteligentes, onde um dos focos é o da mobilidade urbana inteligente (uso racional dos diversos meios de transporte, integrando-os e adaptando-os à demanda). O Brasil já tem tecnologias e infraestrutura para iniciar a implementação de uma agenda nesse sentido. Entretanto, não exploramos ainda todas as possibilidades em relação ao uso de AM e sua relação com comunicação interveicular, entre veículos e a infraestrutura viária, e dentre a própria infraestrutura viária.

Nesse contexto, diversos pesquisadores apontam para um cenário onde a internet chegará aos carros e substituirá, ao menos em parte, a internet como a conhecemos hoje. Por exemplo, em *Reinventing the Automobile*, Mitchell et al. (2010) afirmam que a chamada *mobility internet* possibilitará aos veículos o mesmo a internet atualmente possibilita aos computadores: troca de uma enorme quantidade de informação georreferenciada, em tempo real, o que permitirá integrar veículos à internet das coisas (IoT). Isso potencialmente influenciará o modo como se gerenciam e otimizam as viagens em uma rede viária. Entretanto, muitos desses serviços têm objetivos conflitantes quando estudados em nível de componente do sistema. Por exemplo, sabe-se que a simples difusão do *mesmo tipo de informação* para cada motorista pode ter consequências negativas (Wahle et al., 2000; Wahle; Bazzan; Klügl, 2002). Um segundo exemplo refere-se ao gerenciamento de semáforos (ainda que inteligentes), onde o desempenho de cada um está fortemente atrelado a padrões que estão ocorrendo em cruzamentos adjacentes.

Embora o cenário imaginado por Mitchell e colegas não seja ainda visível no mundo real, ele já é trabalhado em laboratórios de pesquisa como o MAS-Lab, o laboratório do grupo de sistemas multiagente do Instituto de Informática da Universidade Federal do Rio Grande do Sul (UFRGS). Nele, há mais de duas décadas, a autora propõe, desenvolve e aplica técnicas de IA a diversos problemas ligados à questão da mobilidade urbana como um dos fatores subjacentes ao desenvolvimento de políticas públicas que levem a cidades mais inteligentes. Citando Martin Wachs (IEEE 2011 *forum on integrated sustainable transportation systems*), “*mobility is perhaps the single greatest global force in the quest for equality of opportunity because it plays a role in offering improved access to other services*”,

Dessa forma, o objetivo deste texto é discutir os diversos trabalhos realizados pela autora no decorrer de sua trajetória. No presente artigo, o foco é no uso de aprendizado por reforço, uma das vertentes do AM. Tais trabalhos objetivam tanto otimizar a oferta (por exemplo, com controle inteligente de semáforos) quanto distribuir a demanda (com a disseminação não trivial de informações e recomendações aos viajantes).

A sequência do texto joga luz em alguns aspectos a respeito de como a IA vem contribuindo e como seu papel se torna cada vez mais decisivo. Assim, um verdadeiro trânsito inteligente resultará de indivíduos, semáforos e veículos conectados e trabalhando em conjunto. Nessa visão, semáforos inteligentes são alimentados com informação a respeito do estado da rede de tráfego, sobre os semáforos vizinhos, eventos imprevistos, e outras informações.

Referencial teórico e estado do conhecimento

O problema de como se deslocar de A até B de maneira eficiente parece ficar cada vez mais complexo e está entre as principais preocupações do cidadão urbano típico. Como chegamos a esse cenário? Colocando de forma simples: toda vez que a demanda excede a oferta, ocorrem congestionamentos.

Pelo caráter multidisciplinar desses tópicos, nesta seção são apresentados de forma sucinta alguns conceitos fundamentais que embasam os trabalhos discutidos na seção que segue. Uma explicação mais detalhada sobre sistemas de transporte e simulação de tráfego pode ser encontrada em Bazzan e Klügl (2007, 2013a,b), Klügl e Bazzan (2012); uma visão sobre a tecnologia baseada em agentes e sobre IA distribuída pode ser consultada em Bazzan (2010) e Bittencourt (2001). As próximas que seguem abordam questões diretamente relacionadas ao objetivo do usuário do sistema viário: como ir de A até B, o que envolve mobilidade urbana, sistemas inteligentes de transporte, e o problema de como alocar viagens à infraestrutura existente, bem como a ideia geral por trás de AR. Na sequência, as subseções “Trabalhos relacionados: controle semafórico” e “Trabalhos relacionados: escolha de rota” elencam o estado-do-conhecimento nas áreas de controle semafórico e navegação guiada, respectivamente.

Sistemas inteligentes de transporte

Os profissionais e técnicos que atuam na área de engenharia de transportes e tráfego há muito tempo já trabalham com ferramentas computacionais que permitem estimar demandas e adequar a oferta de infraestrutura. A área de Intelligent Transportation Systems (ITS) tem um caráter multidisciplinar e surgiu justamente para, entre outros objetivos, fomentar o uso de novas tecnologias. Entre essas, nos últimos anos, a IA tem revolucionado a área de ITS, permitindo aumentar o desempenho de processos de otimização e controle.

São objetivos de ITS desenvolver sistemas de controle disponíveis, corretos, seguros, escaláveis, persistentes e ubíquos. Entretanto, sistemas de controle de tráfego não podem, sozinhos, resolver os problemas acima mencionados.

O problema da alocação de tráfego

Em uma rede de transporte tem-se uma série de pares Origem-Destino (OD) e diversos caminhos conectando cada um desses pares. O problema da alocação do tráfego (Traffic Assignment Problem – TAP) consiste em alocar as viagens na rede urbana de forma ótima, dadas restrições de capacidade, entre

outras. Em geral cada usuário conhece o melhor caminho entre um par OD, assumindo-se que os caminhos estejam livres de congestionamento. Em situações corriqueiras como hora de pico, o padrão de tráfego muda e rotas que antes eram não ótimas podem passar a ser alternativas atraentes. Os motoristas familiarizados com as condições da rede, como *commuters*, tendem a realizar um processo individual de otimização baseado na sua própria experiência. Em uma situação na qual cada usuário encontrou a rota que tem menor tempo de viagem, nenhum tem incentivo para mudar de rota, o que consiste no estado denominado equilíbrio do usuário (ou de Nash) formulado por Wardrop (1952): nenhum usuário pode melhorar seu desempenho mudando de rota, o que equivale a dizer que todas as rotas têm custo igual (Wardrop's First Principle). Esse equilíbrio é mantido, assumindo-se que nem a demanda e nem a rede mude.

Aprendizado por reforço

Técnicas de AM têm encontrado mais e mais aplicações no contexto de transportes. Em particular, Aprendizado por Reforço (AR) é uma das técnicas mais utilizadas pois permite que as diversas classes de agentes (por exemplo, semáforos, veículos) aprendam a se adaptar ao estado do tráfego mediante a construção de um modelo que diz ao agente que ação realizar em cada estado observado. Dessa forma, o projetista do sistema não precisa dotar o agente de modelos que exigem conhecimento do domínio e/ou instâncias de treinamento que são de difícil obtenção. Há duas grandes variantes das técnicas de AR: baseadas em modelos e independentes de modelos. No primeiro caso, o agente é dotado de um modelo que lhe diz como o ambiente deve se comportar, e qual recompensa esperar. No segundo caso, o agente tem que aprender tal modelo.

No campo do AR independente de modelo, um dos algoritmos mais populares é o chamado Q-learning (Watkins; Dayan, 1992). Esse algoritmo calcula uma aproximação do valor do par estado-ação, denotado $Q(s,a)$, ou seja, o valor da ação a quando o agente a realiza no estado s . No limite de um conjunto infinito de observações de transições de estados, a convergência (para o valor ótimo Q^*) é garantida, desde que todos os pares estado-ação (s,a) sejam visitados infinitamente e frequentemente e que algumas restrições em relação à taxa de aprendizado sejam obedecidas.

Tais garantias valem para o caso de um único aprendiz agindo no ambiente. Entretanto, quando tal aprendiz é colocado em um cenário multiagente, ele não estará mais aprendendo a partir de exemplos vindos de um conjunto fixo. Logo, tal aprendizado deixa de ser trivial.

Diversas soluções para esse problema já foram propostas para o caso monoagente e a “solução” mais popular tem sido a de fazer que os agentes aprendam de forma individual ou independente, isto é, desconsiderando a presença dos demais agentes. Essa abordagem é simples, mas é preciso ter em conta que pode ser ineficiente ou até mesmo ineficaz.

Trabalhos relacionados: controle semafórico

Algoritmos e técnicas de controle semafórico existem há várias décadas e derivam principalmente de técnicas de pesquisa operacional e da área de controle. Aqui são estudados controle em tempo real, bem como otimização *offline*. Mais recentemente, técnicas de IA, IA distribuída e de sistemas multiagentes têm sido empregadas.

Uma abordagem clássica é a sincronização dos semáforos (em modo centralizado) de forma que os veículos possam atravessar uma via arterial em um sentido, com uma velocidade específica, sem paradas (a chamada “onda verde”). Os métodos mais conhecidos aparecem em *softwares*, em geral comerciais, como Transyt (Robertson, 1969), Scoot (Hunt et al., 1981), Scats (Lowrie, 1982), e, mais recentemente, TUC (Diakaki et al., 2003).

O problema do controle semafórico também pode ser abordado do ponto de vista de AR. Na maioria dos trabalhos baseados em AR, o aprendizado é utilizado pelos semáforos a fim de aprender uma política que mapeia os estados (normalmente as filas nas interseções) para ações. Em razão do número de trabalhos que empregam AR no controle semafórico, sugere-se consultar os *surveys* de Bazzan (2009); Mannion; Duggan; Howley (2016); Wei et al. (2019); Yau et al. (2017).

Poucos trabalhos envolvem AR, tanto no controle semafórico quanto nos motoristas, como nosso trabalho apresentado adiante. De fato, essa integração, tão óbvia quanto importante, tem recebido pouca atenção na literatura. No trabalho de Wiering (2000), motoristas e semáforos aprendem simultaneamente. Os controladores semafóricos obtêm informações específicas sobre as rotas dos motoristas (por exemplo, o destino) para calcular uma estimativa do tempo de espera, o que pode ser uma suposição muito forte. Além disso, o modelo subjacente não é totalmente microscópico. O trabalho de Taale, van Kampen e Hoogendoorn (2015) não utiliza AR, mas sim uma estratégia baseada em *back-pressure* para integrar semáforos e influenciar ou substituir a escolha da rota. A abordagem proposta foi testada utilizando apenas o modelo de simulação macroscópico.

Trabalhos relacionados: escolha de rota

Entender como se comporta o motorista em termos de escolha de rota é fundamental em um sistema de recomendação de rotas. Alguns trabalhos já foram desenvolvidos nessa área, como Adler e Blue (1998); Ben-Akiva; Palma; Kaysi (1991); Bonsall (1992); Mahmassani e Chen (1991). Porém, nesses, a resposta do usuário a tais sistemas não é considerada. Isso somente é possível quando se passa ao nível de simulação microscópica baseada em agentes.

Conforme mencionado anteriormente, é fundamental balancear os custos global e individuais. Uma forma de fazer isso é pelo controle de tráfego utilizando sistemas de pedágio especialmente constituídos para penalizar vias com maior tráfego visando melhorar a eficiência da rede pela busca de um equilíbrio no que se refere à distribuição do tráfego. Trabalhos nesse sentido incluem: Arnott, Palma e Lindsey (1990); Kobayashi e Do (2005), Buriol et al. (2010), Sharon

et al. (2017). Porém, pedágios são impopulares e, via de regra, injustos. Uma alternativa busca atingir resultados similares pela disseminação de informação aos usuários. Não são muitos os trabalhos que consideram IA nesse contexto. Redes neurais são utilizadas em Dia e Panwai (2014) para prever a escolha de rota dos motoristas. Entretanto, os autores focam apenas no impacto das mensagens e relegam a segundo plano o impacto da distribuição do tráfego e tempo de viagem. Redes neurais também são utilizadas em Barthélemy e Carletti (2017), onde os parâmetros da rede são determinados em uma fase preliminar de treinamento. A saída da rede neural é a ação a ser realizada pelo agente: permanecer ou modificar o trajeto, a fim de evitar vias congestionadas.

O trabalho de Dias et al. (2014) utiliza um algoritmo da colônia de formigas. A diferença é que, ao invés de usar o feromônio para atrair as formigas, ele inverte esse efeito, repelindo-as. A abordagem proposta por Claes, Holvoet e Weyns (2011) também se baseia em colônia de formigas, combinada com previsão de tráfego na rede viária. No entanto, aqui também os agentes possuem informações centralizadas.

Dessa forma, nota-se que existem claras lacunas na literatura. A próxima seção apresenta métodos propostos no sentido de endereçar, ao menos parcialmente, tais lacunas.

Métodos e resultados

Esta seção está organizada da seguinte forma. As duas primeiras subseções descrevem os principais métodos propostos que lidam, respectivamente, com controle semafórico inteligente e com navegação guiada e escolha de rotas pelos usuários da rede viária (motoristas). A maior parte desses métodos utiliza AR. Uma vez que, no mundo real, tanto os controladores semafóricos quanto os motoristas aprendem simultaneamente, isto leva ao problema de coaprendizado. A subseção posterior discute trabalhos que abordam esse cenário.

Controle semafórico inteligente usando aprendizado de máquina e outras técnicas de IA

As abordagens clássicas descritas na subseção anterior “Trabalhos relacionados: controle semafórico” (Transyt etc.) apresentam algumas desvantagens. Na tentativa de saná-las, a seguir são descritas as principais abordagens propostas e desenvolvidas pela autora.

• Coordenação de semáforos via Teoria dos Jogos

Na sua tese de doutorado, a autora propôs a primeira abordagem para formação de ondas verdes onde cada semáforo é modelado como um agente que aprende. Esse trabalho foi posteriormente publicado em Bazzan (2005). Cada semáforo possui planos predefinidos para sincronização/coordenação com agentes adjacentes em diferentes sentidos de acordo com a situação do tráfego. Essa abordagem utiliza técnicas de AR e teoria dos jogos evolucionária e foi testada em uma via arterial, obtendo desempenho melhor que uma abordagem centralizada clássica.

• Abordagem de AR baseado em modelo

Ao lidar com ambientes não estacionários, onde o fluxo de veículos não é constante, tanto as abordagens de AR independentes de modelo quanto as baseadas em modelo apresentam problemas. Especificamente, quando o ambiente muda, ambas precisam reaprender a partir do zero, uma vez que a política calculada para um dado ambiente deixa de ser válida após uma mudança na dinâmica do fluxo de veículos. Isso faz que os algoritmos de AR apresentem quedas de desempenho durante a fase de reaprendizado, e também forçam o algoritmo a reaprender políticas mesmo para dinâmicas que já tenham sido previamente experimentadas. Além disso, os métodos de AR assumem a existência de um número fixo de modelos de comportamento do ambiente. Uma vez que essa suposição nem sempre é realística, uma opção é a construção incremental de modelos para descrever o comportamento da rede viária.

Um método com essa característica foi proposto em Silva et al. (2006) onde controlador semafórico é capaz de particionar o conhecimento de forma que cada modelo automaticamente trata de um tipo de fluxo. A cada modelo, associa-se uma política ótima, que consiste em um mapeamento entre condições de tráfego e o correspondente plano semafórico a ser escolhido.

Além do melhor desempenho obtido quando comparado com métodos livres de modelo, este método é ainda hoje amplamente citado, tendo aberto novos rumos na pesquisa de métodos de AR baseados em modelo.

• Abordagens baseadas em controle hierárquico

A literatura indica que redes viárias com muitos controladores semafóricos apresentam um desafio para métodos de AR. Para lidar com esse problema de redes viárias de larga escala, em Bazzan, Oliveira e Silva (2010) foi proposta uma abordagem que explora a organização dos agentes em grupos, cada qual sob supervisão de um agente-gerente. Esse gerente faz um controle de várias intersecções, a fim de implementar ações conjuntas e não apenas locais. Essas, por sua vez, tentam balancear as ações recomendadas pelo gerente com as ações que levam ao desempenho ótimo do ponto de vista local. Nessa mesma direção, Abdoos, Mozayani e Bazzan (2013) propõem o uso de um sistema multiagente holônico para modelar uma rede viária particionada em regiões (holons). O diferencial desse método foi a extensão do método Q-learning para o nível de região. Em ambos os casos, a principal vantagem obtida foi uma aceleração significativa no aprendizado, o que equivale a dizer que os semáforos se adaptam muito mais rapidamente às novas condições de fluxo.

Ajudando o motorista a navegar

Enquanto a seção anterior tratou de controle semafórico, esta subseção aborda o lado da demanda, focando naquilo que permeia a questão básica de todos os usuários da rede viária: como ir de A até B de forma ótima (menor tempo, menor custo etc.)? Esse problema tem tido diversas abordagens, segundo as possíveis tecnologias disponíveis na época em questão. No caso da pesquisa mais

recente, o foco é em: disseminação de informação, comunicação veicular, como aprender a escolher rotas, efeito de mudanças de comportamento da parte dos motoristas na presença de informação, e como disseminar informação de modo a garantir um determinado nível de desempenho do sistema.

Para atingir tais objetivos, a autora tem proposto diversos métodos, alguns pioneiros ao abordar a disseminação de informação via dispositivos móveis quando o *smartphone* não existia como o conhecemos hoje. Outros métodos – descritos nas subseções seguintes – envolvem e envolveram: teoria de jogos, comunicação C2C, escolha de rota via AR, e efeito de recomendação de rotas para alinhamento do equilíbrio do usuário ao ótimo do sistema, em oposição à imposição de pedágios.

- **Ser humano no Loop: modelando a escolha de rotas através do uso de teoria dos jogos**

A disseminação de novas tecnologias visando informar o motorista traz consigo a necessidade de considerar o ser humano no laço de controle e alocação de tráfego, o chamado *human-in-the-loop*. No passado, essa questão havia sido de certa forma deixada em segundo plano pela impossibilidade de tratá-la computacionalmente. Com o aumento do poder computacional dos processadores, o advento da modelagem baseada em agentes, e também de diversos projetos multidisciplinares, diversas tentativas de modelar o problema de escolha de rotas foram feitas, incluindo as realizadas pela autora, que levam em conta a heterogeneidade de tais decisões; afinal cada agente tem sua forma particular de tomar decisão. Para tanto foi necessário desenvolver alguns modelos referentes ao comportamento do motorista, como o uso de um modelo do estado mental (por exemplo as crenças, intenções e desejos do motorista) usando lógica BDI (Beliefs, Desires, Intentions). Podem-se considerar como componentes desses modelos: desejos relacionados a maximizar o tempo de tráfego, e crenças acerca do estado e custo de cada rota ou parte da malha viária usada pelo agente. Uma aplicação pode ser encontrada em Bazzan, Wahle e Klügl (1999), que trouxe o diferencial de poder modelar agentes de forma individual, considerando, portanto, a heterogeneidade da população.

Um dos problemas encontrados foi o fato de que ainda não se têm modelos precisos de como funciona o processo do raciocínio humano subjacente à decisão sobre rotas. A fim de investigar esse processo, uma ferramenta útil é a teoria de jogos experimental. Enquanto a teoria de jogos clássica fornece várias ferramentas para modelar *congestion games*, em experimentos com humanos, é possível observar se e como esses desviam dos resultados da teoria clássica.

Em um projeto com participação da autora, coordenado pelos professores M. Schreckenberg e R. Selten (projeto Survive), projetamos experimentos onde humanos escolhem entre duas rotas de forma iterativa, a partir de informações sobre as escolhas anteriores. O objetivo principal foi estudar o efeito da disseminação de tipos de informações aos viajantes através de dispositivos móveis.

Ressalta-se que essa ideia original antecedeu o efetivo emprego desses dispositivos – que só viriam ao mercado em 2008 – e dos aplicativos de trânsito, somente disponibilizados anos depois. Os dados coletados nesse experimento foram a base para a formulação de heurísticas para escolha iterativa de rotas publicadas em Klügl e Bazzan (2004a,b), onde uma forma simples de AR simulava as escolhas de fato realizadas pelos humanos.

Esse trabalho teve implicações importantes, pois, até alguns anos atrás, os meios típicos de difusão de informação sobre o estado do trânsito eram rádio e TV ou painéis de mensagens variáveis. Não apenas não eram feitas recomendações explícitas de rota, como também essas informações não atingiam a totalidade dos usuários. Entretanto, atualmente são empregados outros meios baseados em comunicação, internet e geoposicionamento, mediante serviços como Waze e similares. Esses, de posse da localização de uma massa significativa de usuários do serviço, recomendam uma rota para o usuário. Um problema aqui é que a recomendação, se for a mesma para todos os usuários, pode levar ao colapso da rota recomendada caso a maioria dos usuários siga a recomendação. Simulações desse tipo de cenário aparecem em Bazzan, Fehler e Klügl (2006); Bazzan e Klügl (2005); Klügl e Bazzan (2004a).

Essas são questões conhecidas na teoria de jogos. Em problemas ligados a *minority games*, como em Bazzan e Klügl (2005), é sabido que em sistemas onde cada participante tenta otimizar seu desempenho de forma individual e gulosa, o desempenho global é ruim. No caso de redes de transportes, assumindo-se que veículos tentem evitar uma rota A optando por uma rota B, essa última terá perda drástica de desempenho, podendo haver inclusive situações de oscilação e deterioração para todos os participantes, como em Wahle et al. (2000). Especificamente, essa é uma questão que fundamenta o chamado paradoxo de Braess, originalmente apresentado em Braess (1968), que representa um fenômeno contra intuitivo: em uma malha viária, quando é adicionada uma nova rota, de baixo custo, conectando dois pontos, é possível que haja não redução no tempo de percurso, mas sim um aumento. Isso ocorre pois cada decisão dos motoristas (baseada em sua estimativa de custo) ignora os efeitos das decisões de outros motoristas sobre o cenário global. Ou seja, os motoristas, ao tentarem reduzir seus tempos de tráfego individualmente, de maneira gulosa, acabam aumentando o tempo de tráfego global. O objetivo global se opõe aos objetivos individuais de alguns motoristas.

Conforme mostrado em Bazzan e Klügl (2005), o uso de AR faz que os motoristas se adaptem e aprendam a evitar rotas ineficazes, melhorando o desempenho do sistema como um todo.

- **Aprendendo a escolher rotas**

Nos artigos mencionados na subseção anterior, AR foi usado de forma simplificada, sem considerar fatores como mudança de rota durante as viagens, o arrependimento dos agentes por escolhas ineficazes, a granularidade da simu-

lação (se microscópica ou macroscópica), e a busca pelo desempenho ótimo do sistema.

Esses pontos foram abordados no MASLab por métodos como Ramos e Grunitzki (2015) (uso do formalismo *learning automata*); em Bazzan e Grunitzki (2016) (modelagem como jogo estocástico, onde o agente pode alterar sua rota a cada cruzamento, a fim de navegar até seu destino; em Ramos, Bazzan e Silva (2018) (escolha de rotas considerando o arrependimento em escolhas passadas); e em Cagara, Scheuermann e Bazzan (2015); Bazzan (2019). No caso desses dois últimos trabalhos, o objetivo é alinhar o ótimo do sistema ao equilíbrio do usuário. Esse problema, conforme mencionado na subseção anterior, está ligado ao paradoxo de Braess, onde a adição de recursos aos sistemas degrada seu desempenho. Para mitigar esse problema, Cagara, Scheuermann e Bazzan (2015) propuseram o uso de um algoritmo genético distribuído. No modelo de “ilhas”, cada veículo, pela comunicação veicular, troca soluções com veículos que encontra na via. No caso de Bazzan (2019), um algoritmo genético troca informações com um processo de AR que é executado no nível dos agentes motorista/veículo. Essa sinergia permite que ambos enviem as soluções, um do outro, não apenas acelerando a convergência dos aprendizados, como também levando a soluções mais eficientes em nível do sistema (ótimo global).

Integrando controle inteligente e navegação Guiada

As subseções “Controle semafórico inteligente usando aprendizado de máquina e outras técnicas de IA” e “Ajudando o motorista a navegar” apresentaram diversos métodos propostos e desenvolvidos no MASLab para abordar questões de controle inteligente de semáforos e escolha de rotas, respectivamente. Entretanto, no mundo real, essas duas tarefas ocorrem de forma não apenas simultânea, como também altamente acoplada. Claramente, o aprendizado dos controladores semafóricos afeta o aprendizado dos motoristas e vice-versa. Assim sendo, é importante considerar a adaptação simultânea das duas classes de agentes.

Em Bazzan et al. (2008) foram lançadas as bases para o AR nas duas classes. Porém, o modelo de simulação não era plenamente microscópico. Dessa forma, em Lemos e Bazzan (2019) e Lemos, Bazzan e Pasin (2018) a tarefa de aprendizado simultâneo das duas classes de agentes foi estendida para o ambiente de simulação microscópica, o que levou a vários desafios. Em primeiro lugar, o problema de termos duas classes de agentes representando a oferta e a demanda adaptando-se simultaneamente torna o problema mais complexo, computacionalmente falando, pois diversas garantias de convergência perdem sua validade. Em segundo lugar, a tarefa de aprendizado torna-se mais complexa pois as ações dos agentes são altamente acopladas. Um outro desafio é o fato de que a natureza dessas tarefas de aprendizado é diferente; o objetivo de um motorista é minimizar seu tempo de viagem individual; o objetivo de um controlador semafórico é reduzir as filas localmente.

Assim sendo, em Lemos e Bazzan (2019) e Lemos, Bazzan e Pasin (2018) foi proposta uma abordagem baseada em jogos repetidos (para a classe motorista) e jogos estocásticos (para os semáforos). Enquanto a primeira tarefa de aprendizado baseia-se em episódios não sincronizados (em nível dos agentes motoristas), os semáforos têm horizonte infinito de aprendizado. Por fim, o modelo de simulação é microscópico, um aspecto inovador desse trabalho.

Os experimentos foram realizados utilizando um simulador microscópico e uma rede viária em grade com 32 semáforos. Foi mostrado que o coaprendizado apresentou melhores resultados em termos do tamanho das filas de veículos nos semáforos, bem como que os tempos de viagem foram reduzidos.

Conclusão

O tema mobilidade urbana é uma das prioridades do gestor público. Neste trabalho foram elencados alguns dentre os diversos aspectos da questão, focando em controle semaforico e navegação guiada. Foi mostrado que é possível melhorar a eficiência global do sistema por meio de novos métodos e tecnologias ligadas à inteligência artificial em geral, e, em particular, ao aprendizado de máquina e aprendizado por reforço.

Foram apresentados métodos baseados em aprendizado de máquina, que evidenciam os ganhos em termos de tempo de viagem e eficiência do sistema como um todo. Um foco importante desses trabalhos tem sido o estudo do impacto da disseminação de informação ao usuário. Além disso, o tema da disseminação de informação para navegação guiada começa a ser investigado de forma integrada ao uso de aprendizado por reforço para controle semaforico. Tais tecnologias, quando aliadas aos trabalhos em desenvolvimento pela comunidade de engenharia de tráfego, tendem a ajudar o gestor público a melhorar a qualidade do serviço provido ao cidadão, levando a uma melhora da qualidade de vida nas metrópoles do país.

Referências

- ABDOOS, M.; MOZAYANI, N.; BAZZAN, A. L. Holonic multi-agent system for traffic signals control. *Engineering Applications of Artificial Intelligence*, v.26, n.5-6, p.1575-87, 2013.
- ADLER, J. L.; BLUE, V. J. Toward the Design of Intelligent Traveller Information Systems. *Transportation Research Part C*, v.6, p.157-72, 1998.
- ARNOTT, R.; PALMA, A. de; LINDSEY, R. Departure Time and Route Choice for the Morning Commute. *Transportation Research B*, v.24, p.209-28, 1990.
- BARTHÉLEMY, J.; CARLETTI, T. A dynamic behavioural traffic assignment model with strategic agents. *Transportation Research Part C: Emerging Technologies*, v.85, p.23-46, 2017.
- BAZZAN, A. L. C. A Distributed Approach for Coordination of Traffic Signal Agents. *Autonomous Agents and Multiagent Systems*, v.10, n.1, p.131-64, March 2005.

BAZZAN, A. L. C. Opportunities for Multiagent Systems and Multiagent Reinforcement Learning in Traffic Control. *Autonomous Agents and Multiagent Systems*, v.18, n.3, p.342-75, June 2009.

_____. IA Multiagente: mais inteligência, mais desafios. In: MEIRA JUNIOR, W.; CARVALHO, A. C. P. L. F. de (Ed.). *Atualizações em informática 2010*. Rio de Janeiro: PUC-Rio, 2010. p.111-59.

_____. Aligning individual and collective welfare in complex socio-technical systems by combining metaheuristics and reinforcement learning. *Eng. Appl. of AI*, v.79, p.23-33, 2019.

BAZZAN, A. L. C.; FEHLER, M.; KLÜGL, F. Learning To Coordinate In a Network Of Social Drivers: the role of information. In: INTERNATIONAL WORKSHOP ON LEARNING AND ADAPTATION IN MAS (LAMAS 2005), 2006. *Proceedings*. . . S.l.: s.n., 2006. n.3898, p.115-28. (Lecture Notes in Artificial Intelligence).

BAZZAN, A. L. C.; GRUNITZKI, R. A multiagent reinforcement learning approach to en-route trip building. In: INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS (IJCNN), 2016., 2016. *Anais*. . . s.l.: s.n., 2016. p.5288-95.

BAZZAN, A. L. C.; KLÜGL, F. Case Studies on the Braess Paradox: simulating route recommendation and learning in abstract and microscopic models. *Transportation Research C*, v.13, n.4, p.299-319, August 2005.

_____. Sistemas Inteligentes de Transporte e Tráfego: uma abordagem de tecnologia da informação. In: KOWALTOWSKI, T.; BREITMAN, K. K. (Ed.). *Anais das Jornadas de Atualização em Informática*. s.l.: SBC, 2007.

_____. *Introduction to Intelligent Systems in Traffic and Transportation*. s.l.: Morgan and Claypool, 2013a. 1-137p. n.3. (Synthesis Lectures on Artificial Intelligence and Machine Learning, v.7).

_____. A review on agent-based technology for traffic and transportation. *The Knowledge Engineering Review*, v.29, n.3, p.375-403, 2013b.

BAZZAN, A. L. C.; OLIVEIRA, D. de; KLÜGL, F.; NAGEL, K. Adapt or not to adapt – Consequences of adapting driver and traffic light agents. In: TUYLS, K.; NOWÉ, A.; GUESSOUM, Z.; KUDENKO, D. (ed.). *Adaptive agents and multi-agent systems III*. [S.l.]: Springer-Verlag, 2008, v.4865, p.1-14. (Lecture Notes in Artificial Intelligence).

BAZZAN, A. L. C.; OLIVEIRA, D. de; SILVA, B. C. da. Learning in Groups of Traffic Signals. *Eng. Applications of Art. Intelligence*, v.23, p.560-8, 2010.

BAZZAN, A. L. C.; WAHLE, J.; KLÜGL, F. Agents in Traffic Modelling - from Reactive to Social Behavior. In: ADVANCES IN ARTIFICIAL INTELLIGENCE, 1999, Berlin/Heidelberg. *Anais*...Springer, 1999. n.1701, p.303-6. (Lecture Notes in Artificial Intelligence). Extended version appeared in Proc. of the U.K. Special Interest Group on Multi-Agent Systems (UKMAS), Bristol, UK.

BEN-AKIVA, M.; PALMA, A. de; KAYSI, I. Dynamic Network Models and Driver Information Systems. *Transp. Res. A*, v.25, n.5, p.251-66, 1991.

BITTENCOURT, G. *Inteligência Artificial: ferramentas e teorias*. 2.ed. Florianópolis: Editora da UFSC, 2001.

- BONSALL, P. W. The Influence of Route Guidance Advice on Route Choice in Urban Networks. *Transportation*, v.19, n.1, 1992.
- BRAESS, D. Über ein Paradoxon aus der Verkehrsplanung. *Unternehmensforschung*, v.12, p.258, 1968.
- BURIOL, L. S. et al. A biased random-key genetic algorithm for road congestion minimization. *Optimization Letters*, v.4, p.619-33, 2010.
- CAGARA, D.; SCHEUERMANN, B.; BAZZAN, A. L. Traffic optimization on Islands. In: IEEE VEHICULAR NETWORKING CONFERENCE (VNC 2015), 7., 2015, Kyoto, Japan. *Anais. . . IEEE*, 2015. p.175-82.
- CLAES, R.; HOLVOET, T.; WEYNS, D. A decentralized approach for anticipatory vehicle routing using delegate multiagent systems. *IEEE Transactions on Intelligent Transportation Systems*, v.12, n.2, p.364-73, March 2011.
- DIA, H.; PANWAI, S. *Intelligent Transport Systems: neural agent (neugent) models of driver behaviour*. s.l.: LAP Lambert Academic Publishing, 2014.
- DIAKAKI, C. et al. Extensions and New Applications of the Traffic Signal Control Strategy TUC. In: ANNUAL MEETING OF THE TRANSPORTATION RESEARCH BOARD, 82., 2003. *Proceedings. . . s.l.: s.n.*, 2003. p.12-16.
- DIAS, J. C. et al. An Inverted Ant Colony Optimization approach to traffic. *Engineering Applications of Artificial Intelligence*, v.36, n.0, p.122-33, 2014.
- HUNT, P. B. et al. *A Traffic Responsive Method of Coordinating Signals*. Berkshire: Transport and Road Research Laboratory, 1981. TRRL Lab. Report. (1014).
- KLÜGL, F.; BAZZAN, A. L. C. Route Decision Behaviour in a Commuting Scenario. *Journal of Artificial Societies and Social Simulation*, v.7, n.1, 2004a.
- _____. Simulation Studies on Adaptive Route Decision and the Influence of Information on Commuter Scenarios. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, v.8, n.4, p.223-32, October/December 2004b.
- _____. Agent-based modeling and simulation. *AI Magazine*, v.33, n.3, p.29-40, 2012.
- KOBAYASHI, K.; DO, M. The Informational Impacts of Congestion Tolls upon Route Traffic Demands. *Transportation Research A*, v.39, n.7-9, p.651-670, August-November 2005.
- LEMOS, L. L.; BAZZAN, A. L. C. Combining adaptation at supply and demand levels in microscopic simulation: a multiagent learning approach. *Transportation Research Procedia*, v.37, p.465-72, 2019. Selected and peer-reviewed from the 21st EURO Working Group on Transportation (EWGT) meeting.
- LEMOS, L. L.; BAZZAN, A. L. C.; PASIN, M. Co-Adaptive Reinforcement Learning in Microscopic Traffic Systems. In: IEEE CONGRESS ON EVOLUTIONARY COMPUTATION, CEC 2018, RIO DE JANEIRO, BRAZIL, JULY 8-13, 2018, 2018., 2018. *Anais. . . s.l.: s.n.*, 2018. p.1-8.
- LOWRIE, P. The Sydney Coordinate Adaptive Traffic System - Principles, Methodology, Algorithms. In: INTERNATIONAL CONFERENCE ON ROAD TRAFFIC SIGNALLING, 1982, Sydney, Australia. *Proceedings. . . s.l.: s.n.*, 1982.

- MAHMASSANI, H. S.; CHEN, P. S. Comparative Assessment of Origin-based and en route Real-time Information under Alternative user Behavior Rules. *Transportation Research Record*, v.1306, p.69-81, 1991.
- MANNION, P.; DUGGAN, J.; HOWLEY, E. An Experimental Review of Reinforcement Learning Algorithms for Adaptive Traffic Signal Control. In: MCCLUSKEY, L. T. et al. (Ed.). *Autonomic Road Transport Support Systems*. s.l.: Springer, 2016. p.47-66.
- MITCHELL, W. J.; BORRONI-BIRD, C. E.; BURNS, L. D. Reinventing the Automobile. Cambridge, MA: MIT Press, 2010.
- RAMOS, G. de. O.; BAZZAN, A. L. C.; SILVA, B. C. da. Analysing the impact of travel information for minimising the regret of route choice. *Transportation Research Part C: Emerging Technologies*, v.88, p.257-71, Mar 2018.
- RAMOS, G. de. O.; GRUNITZKI, R. An Improved Learning Automata Approach for the Route Choice Problem. In: KOCH, F.; MENEGUZZI, F.; LAKKARAJU, K. (Ed.). *Agent Technology for Intelligent Mobile Services and Smart Societies*. s.l.: Springer Berlin Heidelberg, 2015. p.56-67. (Communications in Computer and Information Science, v.498).
- ROBERTSON, D. I. *TRANSYT*: a traffic network study tool. London: Road Res. Lab., 1969. Rep. LR. (253).
- SHARON, G. et al. Real-time Adaptive Tolling Scheme for Optimized Social Welfare in Traffic Networks. In: INTERNATIONAL CONFERENCE ON AUTONOMOUS AGENTS AND MULTIAGENT SYSTEMS (AAMAS 2017), 16., 2017, São Paulo. *Proceedings*. . . IFAAMAS, 2017. p.828-36.
- SILVA, B. C. d. et al. Dealing with Non-Stationary Environments using Context Detection. In: INTERNATIONAL CONFERENCE ON MACHINE LEARNING ICML, 23., 2006. *Proceedings*. . . New York: ACM Press, 2006. p.217-24.
- TAALE, H.; VAN KAMPEN, J.; HOOGENDOORN, S. Integrated signal control and route guidance based on back-pressure principles. *Transportation Research Procedia*, v.10, p.226-35, 2015.
- WAHLE, J.; BAZZAN, A. L. C.; KLÜGL, F. The Impact of Real Time Information in a Two Route Scenario using Agent Based Simulation. *Transportation Research Part C: Emerging Technologies*, v.10, n.5-6, p.73-91, 2002.
- WAHLE, J. et al. Decision Dynamics in a Traffic Scenario. *Physica A*, v.287, n.3-4, p.669-81, 2000.
- WARDROP, J. G. Some theoretical aspects of road traffic research. *Proceedings of the Institution of Civil Engineers*, Part II, v.1, n.36, p.325-62, 1952.
- WATKINS, C. J. C. H.; DAYAN, P. Q-learning. *Machine Learning*, Hingham, v.8, n.3, p.279-92, 1992.
- WEI, H. et al. A Survey on Traffic Signal Control Methods. *CoRR*, v.abs/1904.08117, 2019.
- WIERING, M. Multi-Agent Reinforcement Learning for Traffic Light Control. In: SEVENTEENTH INTERNATIONAL CONFERENCE ON MACHINE LEARNING (ICML 2000), 2000. *Proceedings*. . . s.l.: s.n., 2000. p.1151-8.
- YAU, K.-L. A. et al. A Survey on Reinforcement Learning Models and Algorithms for Traffic Signal Control. *ACM Comput. Surv.*, v.50, n.3, 2017.

RESUMO – A área de sistemas inteligentes de transporte há muito investiga como empregar tecnologias da informação e comunicação a fim de melhorar a eficiência do sistema como um todo. Isso se traduz basicamente em monitorar e gerenciar a oferta (rede viária, semáforos etc.) e a demanda (deslocamentos de pessoas e mercadorias). A esse esforço, mais recentemente, estão sendo adicionadas técnicas de inteligência artificial. Essa tem o potencial de melhorar a utilização da infraestrutura existente, a fim de melhor atender a demanda. Neste trabalho é fornecido um panorama focado especificamente em duas tarefas onde a inteligência artificial tem contribuições relevantes, a saber, controle semafórico e escolha de rotas. Os trabalhos aqui discutidos objetivam otimizar a oferta e/ou distribuir a demanda.

PALAVRAS-CHAVE: Inteligência artificial, Aprendizado de máquina, Aprendizado por reforço, Sistemas inteligentes de transporte, Mobilidade urbana.

ABSTRACT – The field of intelligent transportation systems has long investigated how to employ information and communication technologies to improve the efficiency of the system as a whole. This basically means to monitor and manage both supply (traffic network, traffic signals etc.) and demand (vehicles, people and goods). More recently, artificial intelligence techniques are being added to this effort, as they have the potential to improve the usage of existing infrastructure to meet the corresponding demand. In this paper, an overview is given, focusing specifically on two tasks where artificial intelligence has made relevant contributions, namely, traffic signal controls and route choices. The works discussed here aim at optimize the supply and/or distribute the demand.

KEYWORDS: Artificial intelligence, Machine learning, Reinforcement learning, Intelligent transportation systems, Urban mobility.

Ana L. C. Bazzan é professora do Instituto de Informática da Universidade Federal do Rio Grande do Sul (UFRGS). @ – bazzan@inf.ufrgs.br / <https://orcid.org/0000-0002-2803-9607>.

Recebido em 26.1.2021 e aceito em 18.2.2021.

¹ Universidade Federal do Rio Grande do Sul, Instituto de Informática, Porto Alegre, Rio Grande do Sul, Brasil.