

DOKTORI (PhD) ÉRTEKEZÉS TÉZISEI

Megerősítéses tanulás
alkalmazásai válogatott ipari
optimalizálási feladatokon

Szerző:
KEGYES Tamás

Konzulensek:
Prof. Dr. habil. ABONYI János
Dr. SÜLE Zoltán

Pannon Egyetem
Műszaki Informatikai Kar
Informatikai Tudományok Doktori Iskola

2025

1 Motiváció és célkitűzések

Kutatásom célja új módszerek és algoritmusok kidolgozása, valamint ezek alkalmazhatóságának elemzése az Ipar 4.0 optimalizálási, ütemezési, kockázatkezelési és irányítási problémáiban. Ezen célkitűzés az autonóm döntéstámogató megoldásokra vonatkozó egyre növekvő igényekre reflektál, amely elengedhetetlen az ipari átalakulás folytatásához a fenntartható és környezetbarát működés elérése érdekében.

A megerősítéses tanulás (Reinforcement Learning, RL) jelentős lehetőséget kínál a mesterséges intelligencia (MI) alkalmazások forradalmasítására, mivel új megközelítést biztosít a gépi tanulás (Machine Learning, ML) fejlesztéséhez, amely lehetővé teszi a nagy léptékű problémák hatékony kezelését. Ezek a technikák, a széles körben használt Dolgok Internete (Internet of Things, IoT) eszközökkel együtt, új lehetőségeket nyitottak meg összetett rendszerek optimalizálására, beleértve a logisztika, projekttervezés, ütemezés és egyéb ipari területek optimalizálását. E lehetőségek kiaknázása alapvető előrelépést eredményezhet az Ipar 4.0 átalakulásában. Az integráció vertikálisan és horizontálisan megerősödik a digitális átalakulás során, növelve a rugalmasságot, miközben az emberi irányításra és felügyeletre összpontosít. Továbbá, az integrált eszközök által generált adatok mennyisége exponenciálisan növekszik, ami magasabb szintű autonóm folyamatokat és döntéseket követel meg. A megerősítéses tanulás értékes eszközként szolgálhat az önoptimalizáló és önszervező Ipar 4.0 megoldások fejlesztéséhez. Az ilyen alkalmazások fejlesztésének fő kihívása, hogy számos módszer, technika és paraméter definiálása szükséges. Mivel ezek meghatározása részletes ismereteket igényel az RL algoritmusok természetéről, munkám egyik fő célja az RL módszerek átfogó áttekintésének biztosítása az Ipar 4.0 és az intelligens gyártás szempontjából.

Az emberi tevékenységek által okozott növekvő környezeti nyomás arra készíti a kormányokat és szervezeteket, hogy korlátozzák vagy csökkentsék ökológiai lábnyomukat. Ennek eredményeképpen az Európai Bizottság stratégiai célként határozta meg ezt a kihívást. A fenntartható gazdasági és társadalmi átalakulás érdekében új Körforgásos Gazdaság Akciótervet hirdettek meg, amely tisztább és versenyképesebb Európát céloz. Számos ajánlás és javaslat született arra vonatkozóan, hogyan kellene kezelni ezeket a problémákat, és a kulcskérdés a fogyasztás csökkentése és az újrahasznosítási arány növelése. Ezek az erőfeszítések vezettek a körforgásos gazdaság koncepciójának megalkotásához, amely sokkal összetettebb, mint pusztán

a gyártás optimalizálása, mivel magában foglalja az ellátási lánc, a szétszerelés és az újrahasznosítás lépéseinek optimalizálását is. A körforgásos gazdaság optimalizálási céljai lefedik a visszagyűjtési és újrahasznosítási folyamatokat, amelyek valamivel bonyolultabbak, mivel ezek inkább sztochasztikus és kevésbé szabályozott folyamatok. A legalapvetőbb modellek determinisztikus bemeneteket feltételeztek, de a fejlettebb modellek már kezelni kezdték a valós problémákban megfigyelhető bizonytalanságokat. Többek között az alapanyagok és azok eloszlásai sztochasztikus változókkal írhatók le. Hasonlóképpen, a szétszerelési feladatok, az ezekhez szükséges folyamatidők, valamint az újrahasznosítási lépések keresleti értékei is lehetnek sztochasztikus jellemzők néhány modellben.

Az ellátási lánc optimalizálása és az összeszerelési láncok kiegyensúlyozása már az 1950-es évek óta intenzíven kutatott területek, míg a szétszerelési láncok fókuszált elemzése közel 40 évvel később kezdődött. Kutatásom másik célja az volt, hogy részletesen bemutassam, hogyan lehet hatékonyan megvalósítani a szétszerelési láncok kiegyensúlyozási problémáit.

Szélesebb értelemben hasonló optimalizálási problémák találhatók, amelyek megfogalmazása vegyes-egészértékű (vagy vegyes folytonos-diszkrét) optimalizálási problémákhoz vezet. Munkám következő célja egy új problémakategória azonosítása és leírása volt, amely a korlátokkal rendelkező sztochasztikus gráfbejárési problémákra vonatkozik. Bár a klasszikus legrövidebb útkeresési problémákra hatékony megoldások ismertek, ezek a módszerek nem működnek korlátokkal és sztochasztikus eloszlású távolságokkal. Ez adott motivációt számomra, hogy a megerősítéses tanulás módszeréhez forduljak. A vegyes-egészértékű optimalizálási problémák leggyakrabban használt gradiensalapú megoldásaival szemben egy új diszkretizációs megközelítésre összpontosítottam, amely az állapottér tisztán diszkrét térre való redukálását végzi, és ezáltal egy iteratív optimalizálási stratégián alapuló módszert alkalmaztam.

Azonosítottam egy sokszínű problémakört, amely a korlátozott gráfbejárési problémákhoz tartozik. Bár a legrövidebb útkeresési probléma jól ismert kombinatorikai optimalizálási probléma, paramétereit nehéz pontosan meghatározni. Könnyen belátható, hogy egy teherautó optimális útvonalának problémája gyakorlatilag egy korlátozott legrövidebb útkeresési probléma, ahol a korlátok a sofőr munkaidőkorlátait, a parkolóhelyek elérhetőségét vagy az üzemanyagtartály kapacitási korlátait írják le. Az elektromos szállító járművek napi tervezése egy korlátozott Hamilton-útkeresési probléma, ahol a korlátokat az akkumulátor kapacitásának,

a töltési lehetőségeknek vagy akkumulátorcserének a korlátai határozzák meg. Továbbá, egy termék élettartama utáni összes komponensének szétszerelése szintén egy korlátozott gráfbejárási probléma, ahol egy előfeltétel-gráf írja le a komponensek eltávolítási függőségeit, és a korlátok a párhuzamos munkaállomások használatát is korlátozzák. Végezetül meg kell jegyezni, hogy a valós problémákban ezek a problémák inkább sztochasztikus optimalizálási feladatok, mintsem diszkrét.

A megerősítéses tanulás (RL) kézenfekvő megoldást jelenthet szekvenciális döntéshozatali folyamatokhoz, például lépésről lépésre történő útkereséshez. A célfüggvényt a jutalomfüggvényben kell megvalósítani, valamint a korlátozások megsértését következményeit is. Azonban sztochasztikus hatások vagy mérési bizonytalanság figyelembevételével további nehézségek merülnek fel: néhány folytonos komponens integrálni kell az állapottérbe, amely vegyes diszkrét-folytonos típusú lesz. Vannak sikeres módszerek a vegyes-egészértékű problémák gradiensalapú vagy mély RL módszerekkel történő megoldására. A korlátos kombinatorikai optimalizálási problémák szintén megoldhatók RL technikákkal. Sőt, a sztochasztikus legrövidebb útkeresési problémákhoz is vannak RL megoldások. Ezekhez képest kutatásom célja egy új algoritmus kidolgozása, amely diszkrét lépéseket hajt végre a DBSCAN klaszterezési módszer alkalmazásával, és diszkrét RL módszert használ vegyes folytonos-diszkrét korlátos sztochasztikus gráfbejárási problémákhoz.

2 Új tudományos eredmények

A következő lista három tézisben foglalja össze az új tudományos eredményeimet.

1. **Az Ipar 4.0 alkalmazás megerősítéses tanulási (RL) módszereinek többdimenziós osztályozása elősegíti az aktuális kutatási trendek azonosítását. Ezen felül egy kérdőív alapú útmutató támogatja a megfelelő módszertan követését az RL modell felépítéséhez, és elősegíti a legmegfelelőbb RL-módszer kiválasztását egy adott adott probléma megoldására.**

- 1.1 Szisztematikus irodalomkutatást végeztem, amely során kulcsszavak szabványosítását és három fő dimenzióba történő kategorizálását végeztem el: az alkalmazott alapelvek, az ipari területek, valamint az alkalmazási módszertan matematikai megközelítései szerinti bontásban.
- 1.2 Elemzést készítettem a fő kutatási trendek, valamint a problémátípusok, az alkalmazott RL-módszerek és az ipari területek közötti összefüggések azonosítására.
- 1.3 Ezek alapján létrehoztam egy kérdőíves útmutatót, amely segít a legmegfelelőbb RL-módszer kiválasztásában adott problémák megoldásához.
- 1.4 Az útmutató továbbfejleszhető, hogy egy automatizált RL-módszer-választó alkalmazás alapját képezze.

Kapcsolódó publikációk: [R1].

2. **Az újonnan kifejlesztett, megerősítéses tanuláson alapuló dLOptRL algoritmus képes megoldani a szétszerelési láncok kiegyensúlyozási problémáit, és hatékonyan támogatja a fenntartható szétszerelési láncok optimális működését.**

- 2.1 Elkészítettem a szétszerelési láncok kiegyensúlyozási problémájának általános megfogalmazását a megerősítéses tanulási módszerek számára, beleértve az állapot- és akciótereket, valamint a jutalomfüggvényt.
- 2.2 Kifejlesztettem a DLOPTRL algoritmust, amely a Heurisztikusan Gyorsított Megerősítéses Tanulási módszerek közé tartozik, és hatékonyabban oldja meg a szétszerelési láncok optimalizálási problémáit, mint az eredeti megerősítéses tanulási módszer.

- 2.3 Mivel a legtöbb lineáris és kvadratikus megoldó korlátokkal rendelkezik, a szétszerelési láncok kiegyensúlyozási problémáját újraformáltam egy vegyes-egészértékű kvadratikus optimalizálási problémává, viszont jelentősen kevesebb döntési változóval, hogy referenciamegoldást biztosítsak a DLOPTRL algoritmus teljesítményének értékeléséhez.
- 2.4 Kidolgoztam egy lépésről lépésre megadott eljárást az RL formalizálás előállítására közvetlenül a MIQP megfogalmazásból.
- 2.5 Bemutattam a DLOPTRL algoritmus alkalmazhatóságát kis- és közepes méretű szétszerelési láncok optimalizálási esettanulmányain. Rámutatottam, hogy az algoritmusom valódi alternatívát kínál a klasszikus vegyes-egészértékű kvadratikus programozási megoldásokkal szemben.

Kapcsolódó publikációk: [R2], [R3].

3. **A Q-tömörítéses módszer egy újonnan kifejlesztett, általánosan alkalmazható megerősítés-tanuláson alapuló megközelítés vegyes egészértékű optimalizálási problémák megoldására, amely dinamikus diszkrét reprezentációt használ az állapottérhez számos komplex optimalizálási probléma esetében.**

- 3.1 Meghatároztam a korlátozott sztochasztikus gráfbejárési problémaosztályt, és megmutattam, hogy számos ütemezési és útvonaltervezési probléma ebbe tartozik. Továbbá meghatároztam azokat a pontos módszereket, amelyekkel a problémák átalakíthatók korlátozott sztochasztikus legrövidebb út feladatokká, és ezáltal szabványosíthatók a megfogalmazások.
- 3.2 Kifejlesztettem a Q-tömörítéses módszert, amely dinamikus diszkrét reprezentációt talál egy vegyes-egészértékű optimalizálási problémához a DBSCAN algoritmus használatával, és megerősítéses tanulási módszerrel oldja meg azt.
- 3.3 A Q-tömörítéses módszer javítja a DLOPTRL algoritmus beépített heurisztikáját a művelettér csökkentése érdekében, és egy emberek által is értelmezhető modellt nyújt időigényes előkészítő lépések nélkül.
- 3.4 Végül bemutattam az algoritmus használhatóságát különböző kiválasztott problémákon, és igazoltam előnyét a rácsalapú diszkretizációs módszerrel szemben.

Kapcsolódó publikációk: [R4], [R5].

3 Az új tudományos eredmények felhasználása

Az első tézisemben bemutatott áttekintés és a kérdőíves útmutató támogatja a kutatókat és fejlesztőket a felkészülési idő lerövidítésében azáltal, hogy lehetővé teszi számukra a legrelevánsabb módszerekre való összpontosítást. Ez egy későbbi automatizált módszerválasztási eljárás alapját is képezheti.

A második tézisemben bemutatott DLOPTRL algoritmus egy teljes körű megoldót kínál a szétszerelési láncok kiegyensúlyozási problémáira, amely képes tanulni közvetlenül egy működő alkalmazásból vagy egy szimulált virtuális ikerből.

A harmadik tézisem meghatározza a korlátozott sztochasztikus gráfbejárési problémák osztályát, amely számos fontos optimalizálási problémát foglal magában, és könnyen bővíthető továbbiakkal. Ez lehetővé teszi, hogy egy megoldási módszert könnyedén átvigyünk az egyik problémaosztályból a másikba. A Q-tömörítéses módszer emberek által is értelmezhető megoldást kínál időigényes előzetes lépések nélkül, ezáltal jelentősen csökkentve a megvalósítási időt. Az algoritmus továbbá képes alkalmazkodni a változó környezetekhez a megfigyelési időablakot módosítása által.

Irodalomjegyzék

Magyar Tudományos Művek Tára (MTMT) személyes kutatói oldal:

<https://m2.mtmt.hu/gui2/?type=authors&mode=browse&sel=10081160>

Tézisekhez kapcsolódó publikációk

- [R1] Tamás Kegyes, Zoltán Süle, and János Abonyi. The applicability of reinforcement learning methods in the development of industry 4.0 applications. *Complexity*, 2021:1–31, 2021.
- [R2] Tamás Kegyes, Zoltán Süle, and János Abonyi. Disassembly line optimization with reinforcement learning. *Central European Journal of Operations Research*, pages 1–28, 2024.
- [R3] Tamás Kegyes, Zoltán Süle, and János Abonyi. Incorporation of heuristic search into q-learning. In *Short papers of VOCAL 2022 (9th VOCAL Optimization Conference: Advanced Algorithms)*, Budapest, Hungary, 2022.
- [R4] Tamás Kegyes, Alex Kummer, Zoltán Süle, and János Abonyi. Generally applicable q-table compression method and its application for constrained stochastic graph traversal optimization problems. *Information*, 15(4):193, 2024.
- [R5] Tamás Kegyes, Alex Kummer, Zoltán Süle, and János Abonyi. Improvements of the q-compression method for constrained stochastic graph traversal problems. In *10th VOCAL Optimization Conference: Advanced Algorithms: Conference Program*), Budapest, Hungary, 2024.

További publikációk

- [F1] Tamás Kegyes, Zoltán Süle, and János Abonyi. Machine learning-based decision support framework for cbrn protection. *Heliyon*, 2024.
- [F2] Kegyes Tamás, Süle Zoltán, and Abonyi János. Az információmenedzsment szerepe az abv-védelemben. *Nemzetbiztonsági Szemle*, 11(1):62–77, 2023.
- [F3] Tamás Kegyes, Zoltán Süle, and János Abonyi. Szétszerelési láncok optimálisa megerősítéses tanulással. In *Magyar Operációkutatási Konferencia: Absztraktok könyve*, Budapest, Hungary, 2021.