

Opponensi Vélemény

Nagy László doktorjelölt

Ipar 4.0 és 5.0 megoldások fejlesztése ontológiák alapján - modellezés és optimalizálás

c. PhD értekezéséhez

Pannon Egyetem Vegyészmérnöki- és Anyagtudományok Doktori Iskola

Konzulensek: Prof. Dr. habil. Abonyi János Pannon Egyetem

Dr. Ruppert Tamás Pannon Egyetem

Bíráló: Dr. Gyulai Dávid Western Digital Corporation (HGST Netherlands BV)

1. Témaválasztás, időszerűség

A jelölt a doktori kutatása során gyártó-, ezen belül is szerelőrendszerekkel kapcsolatos gyártásoptimalizálási és modellezési feladatokat vizsgált, és oldott meg korszerű, ontológiák, tudásgráfok és hálózateméleti modelleken alapuló módszerek segítségével. A témaválasztás ipari szempontból megfelelő, hiszen a szerelőrendszerek tipikusan emberi erőforrás-intenzívek, ezáltal különösen rugalmasak a folyamat szempontjából, ez a rugalmasság azonban csak megfelelő tervezés mellett hasznosítható a gyakorlatban, ahol az emberi és a gépi erőforrásoknak összhangban kell lenniük egymással valamint a termelési környezet többi elemével, és magával a folyamattal is. A jelenleg intenzíven zajló ipari digitalizáció és digitális transzformáció a szerelőrendszerek kapcsán is sok új lehetőséget biztosít, ezeket közül a doktorjelölt a rendelkezésre álló adatgyűjtési és folyamat monitorozási lehetőségeket (szenzortechnológia), az ember-robot kollaborációt valamint az emberi erőforrás rugalmas alkalmazását hasznosítja a javasolt modellekben és megoldásokban. A dolgozatban esettanulmányként vizsgált kábelkorbács-szerelési folyamat jól érthető példa, iparilag releváns, valamint nagy mértékben megkönnyíti a dolgozat megértését azáltal, hogy a különböző megoldások mindegyike egy közös területen kerül bemutatásra. Az ipari relevancia mellett fontos kiemelni, hogy a vizsgált példa kellően bonyolult ahhoz, hogy indokolja új modellek és módszerek alkalmazását a felmerülő problémák megoldásához. Az előzők alapján megállapítható tehát, hogy az értekezés témaválasztása korszerű és aktuális, ipari és tudományos is releváns és fontos területet és problémákat vizsgál illetve old meg.

A doktorjelölt a munkája során kreatív módon alkalmazta termelési környezetben a különböző hálózateméleti, modellezési, optimalizálási és elemzési eszközöket, ezeket pedig hatékonyan támogatta olyan számítástudományi és informatikai megoldásokkal mint például a tudásgráfok valamint az ontológiák. A jelölt kellő jártasságot szerzett a szerelőrendszerek adat- és tudásalapú modellezésében, és ezeket modelleket ipari környezetben is hatékonyan képes hasznosítani, például kollaboratív szerelőrendszer tervezésére valamint sorkiegyenlítési feladatok megoldására. Ugyan a kidolgozott új módszerek implementálását egyes esetekben kutatótársaival közösen végezte, de ezt minden esetben a megfelelő helyen jelzi, valamint ez a tudományos eredmények értékéből nem von le.

2. Az értekezés felépítése, eredményei

Az értekezés teljes terjedelme 245 oldal, ebből ~170 a dolgozat törzse valamint az új tudományos eredmények ismertetése, 55 oldal pedig a vizsgált problémákat részletező függelék, a jelölések jegyzéke, valamint az irodalomjegyzék. Érdemben kiemelendő, hogy a dolgozat angolul, megfelelő nyelvezetben, jól érthető fogalmazással íródott, könnyen olvasható és jól érthető. Nyelvtani hibát és elütést keveset tartalmaz (néhány gépelési hiba az új fejezetekben is található), ezek azonban alig észrevehetőek és nem csökkentik a dolgozat és a kutatómunka színvonalát. A dolgozat témamegjelöléssel és problémafelvetéssel kezdődik, amely jelen esetben segítséget nyújt az egyébként meglehetősen tág vizsgált terület, és a dolgozat komplex felépítésének megértéséhez.

A dolgozat fő része ezt követően két fő fejezetre tagolt (az alábbi magyar fordítás talán egyik esetben sem igazán megfelelő): 1. Szemantikus modellezés – ontológiák és tudásgráfok; 2. Hálózattudomány-alapú folyamat optimalizálás – “Fejlett” gyártás-analitika. Ezt követően az első fejezet három, a második pedig négy további alfejezetre tagolt, a különböző problémáknak és megoldásoknak megfelelően. Az új tudományos eredmények a dolgozat törzséhez kapcsolódóan de külön fejezetben kerülnek bemutatásra: a jelölt öt új tudományos állítást fogalmaz meg, ezek helyességét négy saját (témavezetőikkel közös) publikációval igazolja. A negyedik tézispont tartalmaz új eredményeket, ezek publikálása jelenleg is folyamatban van, a tézispont helyességét újonnan ismertetett, részletes teszteredmények igazolják. A függelék a vizsgált rendszer felépítését valamint a folyamat paramétereit mutatja be kellő részletességgel. A bemutatás az adatok mennyisége miatt és esetleg (ipari szempontból) érzékenysége miatt sem lehet teljes körű, ezért teljes mértékben a módszerek az adott példán talán nem reprodukálhatóak, de ez nem von le a tudományos eredményekből és a dolgozat értékéből. A jelölések jegyzéke megfelelő. Az irodalomjegyzék meglehetősen terjedelmes, 368 hivatkozást tartalmaz, ezek többsége folyóirat-publikáció az elmúlt 10-15 évből, vagyis kellően releváns és aktuális. Az irodalomjegyzékben a már említett négy önálló publikáció is szerepel, ezek mindegyike színvonalas fórumokon jelent meg, közülük 3 impakt faktoros folyóirat valamint egy nemzetközi konferenciaközlemény. A hivatkozások száma és relevanciája bőven megfelel az elvárásoknak. A hivatkozások és a jegyzék formátuma egységes, jól strukturált és áttekinthető, az ábrák igényesek és színvonalasak.

3. A fejezetek ismertetése

Fontos kiemelni, hogy a doktorjelölt a dolgozat „munkahelyi vitájára” beadott és általam is véleményezett változatához képest, a korábbi opponensi javaslatok alapján egy javított, átdolgozott kiadást készített. Ez a jelenleg is véleményezett változat tartalmaz olyan új fejezeteket melyek segítik a dolgozat könnyebb megértését, javítja a korábbi esetleges tartalmi, nyelvtani stb. hibákat, illetve az opponensek által korábban hiányolt tartalmi elemeket részletesen taglalja. Ilyen többek között a tudásbázisok és RDBMS-ek összehasonlítása, a sorkiegyenlítési algoritmus eredményeinek részletesebb bemutatása, valamint az általam korábban kifogásolt közösség-detektálási megoldás részletesebb vizsgálata, új metrikák segítségével. Az alábbi ismertető főként az általános, fontosabb kérdéseket és javaslatokat tartalmazza, és figyelembe veszi a doktorjelölt korábbi opponensi véleményemre adott válaszát és javításait a dolgozat új változatában.

A dolgozat **A. fejezete** a témaválasztást mutatja be és indokolja, valamint bemutatja a főbb témák és megoldások definícióját. A dolgozatban új fejezetként jelent meg a korábban hiányolt RDBMS és tudábázis többszempontú összehasonlítása, nem csak tudományos de gyakorlati/felhasználási szempontokból is. Az új fejezet tisztázza a korábban felvetett alkalmazási kérdéseket. A dolgozat vitáján érdemesnek tartom szóban is tárgyalni a gyakorlati alkalmazásokat, azok nehézségeit, előnyeit hátrányait stb.

- Új magyarázó ábra segíti az I5.0 fő elemeit, melyek újonságot jelentenek az I4.0-hoz képest.
- Az A1.4 problémafelvetést a jelölt átdolgozta, így az megfelelően ismerteti a kutatómunka során vizsgált, illetve a dolgozatban bemutatott problémákat. Itt érdemes megemlíteni, hogy a 4. pont inkább a gyártáshoz kapcsolódó *big data* kihívásokra utalhat, a dolgozatban kevés szó esik, és véleményem szerint nem is szerves része a kutatómunkának az ezekkel kapcsolatos problémák megoldása, ez inkább csak határterületként jelenik meg.
- Az angol kifejezések néhány esetben eltérően, nem konzisztensen szerepelnek, pl. *optimisation* ill. *optimization*.

A dolgozat **I. Fő fejezete (Part I)** a szemantikus modellezést és ontológiákat tárgyalja. Sok szó esik a technológiák és standardek használatáról és előnyeiről és hátrányokról, az átdolgozott fejezet a korábbihoz képest jóval érthetőbb és sok korábbi kérdésre választ ad.

- A fejezettel kapcsolatos korábbi hiányérzetemet, miszerint nem került ismertetésre „*vajon miért nem elterjedt még mindig az ipari gyakorlatban (általános, mindennapi alkalmazás), az egyébként technológiai szempontból érett, és tudományosan igazolt hatékonyságú megoldások használata*” új tartalmi elemek ismertetik. Ezt érdemesnek tartom a szóbeli vitán további rövid tárgyalásra.
- Az I.1.1 ábra szemléletesen mutatja be az ontológiák illetve tudásgráfok irodalmának fejlődését: mi lehet az oka annak, hogy előbbi mintha „stagnálna”? Az előnyök és hátrányok összehasonlítása jól ismerteti a két terület közötti különbségeket, de számomra nem egyértelmű hogy ezek alapján ekkor különbségnek kellene mutatkoznia.
- Az ontológiák előnyeinek szerepel, hogy a SPARQL gyorsabb lekérdezéseket tesz lehetővé adott esetekben, ez számomra némileg ellentmond a későbbi pontban megfogalmazott számítási komplexitásoknak.

A dolgozat **II. Fő fejezete (Part II)** szerelőrendszerek hálózat-alapú modellezésével és optimalizálásával foglalkozik, új megoldást javasol sorkiegyenlítési feladat megoldására, közösségek detektálására komplex hálózatokban, valamint ember-robot kollaboratív környezet tervezésére.

- A **II.2 fejezet** egy bonyolult korlátozás-rendszerrel (task-operator-equipment-skill) bíró sorkiegyenlítési feladatra ad heurisztikus megoldást. A bemutatott új tudományos eredményeket a tesztek igazolják, valamint az átdolgozott változat tartalmaz egy érzékenységvizsgálati kísérletsorozatot. Az eredmények alapján a megoldás érzékeny néhány paraméter megváltozására, a kérdésem, hogy ezeket az eredményeket figyelembe véve a gyakorlatban mely megoldást javasolná megvalósításra, illetve mi alapján lehet eldönteni hogy az alternatív megoldások melyike a legmegfelelőbb ipari környezetben?
- A saját eredmények a korábbi változathoz képest hangsúlyosabban jelennek meg a fejezetben, a megoldás skálázhatóságáról továbbra is kevés szó esik. Van erre vonatkozóan valamilyen teszteredmény, esetleg tapasztalat?
- A **II.3 fejezet** egy új, és ötletes megoldást javasol az egyébként nehéz közösség-detektálási feladat megoldására hálózatok esetén. Az új fejezetek kellő részletességgel bevezetik a modularitás

fogalmát, illetve a magyarázó ábrák segítik a könnyebb megértést a szakterületen esetleg kevésbé járta olvasó számára is.

- A II.3.1.2 ismerteti az algoritmus inicializálását, mely jellemzően valamilyen véletlen állapotból indul. A doktorjelölt megemlíti, hogy célszerű lehet többször futtatni az algoritmust különböző, véletlenszerűen generált kezdeti állapotokból. Ilyenkor, adott esetben különböző megoldáshoz jutva, hogyan lehet aggregálni az eredményeket, és javaslatot tenni a végső megoldásra?
- Az új megoldás még nem került publikálásra, ezért az eredmények igazolásánál nagyon fontos hogy több teszt, köztük új, átfogó *benchmark* is igazolja a módszer működését, ez mindenképpen plusz eredmény a fejezetben.
- Az új táblázatos eredményeknél nem világos (II.3.2, II.3.3 táblázatok), hogy mit jelent a C (talán ez a közösségek száma?), és miért van említve m a táblázat szövegében, ugyanis nincs használva. Talán gépelési hiba?
- Ehhez kapcsolódó kérdés, hogy amennyiben $C \rightarrow m$, hogyan lehetséges hogy pl. a második teszt esetén az érték (19) kívül esik a minimum és maximum által megadott (20-50) tartományon?
- Jól értem, hogy a saját megoldás kevésbé érzékeny a paraméteresek változására, a vizsgált egyéb módszerekhez képest? Kérem tisztázni, hogy miért az adott módszereket vizsgálta, illetve valóban ezek-e jelenleg a terület leghatékonyabb algoritmusai.
- A **II. 4 fejezet** több új résszel egészül ki, melyek segítik a fejezet könnyebb megértését, valamint bemutatják a gyakorlati alkalmazások előnyeit illetve korlátait.

4. Összefoglalás

Nagy László komoly kutatómunkát végzett a szerelőrendszerek vizsgálata és modellezése területén, és kreatív, hatékony módon alkalmazott új modellezési eszközöket a hálózatalmélet területéről. Ez igen jó absztrakciós készségre utal, ami kimondottan fontos napjaink gyártórendszereinek modellezése során. A dolgozat egy igen szerteágazó munkát foglal össze szemléletesen és jól érthető módon, a vizsgált közös esettanulmány alapján. **Az elért eredmények tudományos szempontból újak, a módszerek hatékonyak, ezt a publikációk is jól igazolják. A dolgozat formátuma egységes, jól strukturált, az ábrák színvonalasak és nagy mértékben segítik a megértést, főleg az átdolgozott, jelenleg váleményezett változat esetében.**

A dolgozatban bemutatott eredményeket komoly nemzetközi publikációk igazolják, az ezekhez tartozó tézispontok mindegyikét javaslom elfogadásra. A negyedik tézisponthoz nem tartozik publikáció, ezt az új, átfogó számítási eredmények alapján (melyek hamarosan szintén publikálásra kerülnek) szintén elfogadásra javaslom.

Mindezek alapján javaslom az értekezés nyilvános vitára bocsátását, és sikeres védés esetén a PhD cím odaítélését támogatom.

Zandvoort, 2023. május 14.

Gyulai Dávid, PhD