POLINOM IDEJŰ, HEURISZTIKUS OPTIMALIZÁCIÓS MÓDSZEREK ALKALMAZÁSAI A PÉNZÜGYI SZÁMÍTÁSTUDOMÁNYBAN

Ph.D. disszertáció tézisfüzet

Fogarasi Norbert

Témavezető:
Dr. Levendovszky János, D. Sc.
A Magyar Tudományos Akadémia doktora

Híradástechnikai Tanszék
Budapesti Műszaki és Gazdaságtudományi Egyetem

Budapest, 2014

1. Bevezetés

A pénzügyi számítástudomány olyan alkalmazott számítástechnikai tudományos ágazat, amely a pénzügy gyakorlati kérdéseivel foglalkozik. Viszonylag új tudományág, melynek születését Harry Markowitz 1950-es évekbeli munkásságától számítjuk és rendkívül szerteágazó algoritmikus problémákat tartalmaz. Manapság új numerikus módszereket dolgoz ki az optimalizáció, a statisztika és a sztochasztika területén a legkülönfélébb pénzügyi alkalmazások céljából.

Markowitz úgy fogta fel a portfólió-választást, mint a hozam és kockázat egyidejű optimalizációját. Kulcsértékű felfedezése, hogy a diverzifikáció, mint a variancia csökkentésének eszköze, alapvető fontosságú, mivel növeli a befektetés előreláthatóságát, akár a várt hozam csökkenése árán is [32]. Korszakalkotó felfedezéséért Markowitz 1990-ben Nobel-díjat kapott. Számos más problematika is megjelent azonban a pénzügyi számítástudomány területén az 1950-es évek óta, például az algoritmikus kereskedés, a hatékony piaci hipotézisek vizsgálata, az opcióárazás vagy a pénzügyi idősorok előrejelzése.

Az algoritmikus kereskedés területe, az elektronikus platformok alkalmazása a kereskedési megrendelések algoritmikus elindítására, amellyel a megrendelések azon szempontjait határozzuk meg, mint az ütemezés, az ár vagy a mennyiség, az elmúlt évtizedben egyre nagyobb figyelmet kapott. Különösen a magas frekvenciájú kereskedés nyert egyre jelentősebb teret nemzetközi piacokon, ahol számítógépek hoznak algoritmikus döntéseket, gyakorta a másodperc töredéke alatt, valósidejű adatok révén, összehasonlíthatatlanul rövidebb idő alatt, mint ami az embernek szükséges, hogy feldolgozza az észlelt, megfigyelt információkat. 2009 óta az Egyesült Államokban a teljes kereskedési volumennek több mint 70%-át magas frekvenciájú kereskedési algoritmusok adják. Ugyanakkor, a számítógépes hardware technológiák – mint a GPGPU és az FPGA – ugyancsak nagyot fejlődtek, így a hagyományos módszerek, amelyek csak alacsony frekvenciájú kereskedési környezetben voltak hatékonyak, ma már valós időben kivitelezhetők. Másfelől pedig az optimális algoritmikus kereskedés fejlődése mentén jelentkező számos probléma rendkívül összetett, ami a magas frekvenciájú kereskedésben való gyakorlati hasznosítás érdekében gyors approximációs módszereket követel.

A pénzügyi számítástudomány másik fontos területe a Monte Carlo szimulációs technika, amely piaci folyamatokat bizonyos, meghatározott sztochasztikus feltételezésekkel szimulál. Ez pedig folyamatosan növekvő számítógépes kapacitások, erőforrások bevetését feltételezi, miközben ezek optimalizált alkalmazását ütemezés révén oldja meg. Az optimális ütemezés számos problémájának exponenciális vagy NP komplexitású futási idejének bizonyítása tette szükségessé a polinom idejű approximációs technikák alkalmazását.

Ahogy a fentiekben bemutattuk, a szakterület számos alapvető problémája exponenciális komplexitást, vagy NP nehézséget mutat. Ennek az az eredménye, hogy a pénzügyi számítástudomány egyik alapvető célkitűzése olyan numerikus algoritmusok megtalálása, amelyek a gyakorlatban alkalmazható időkeretben megközelítő megoldásokat nyújtanak ezekre a problémákra.

E disszertáció a fenti célkitűzéseket két konkrét problémára alkalmazva vizsgálja:

1) valós historikus adatok alapján olyan ritkás kitöltésű portfóliók megtalálása, amelyek maximalizálják az előrejelezhetőséget; és
2) optimális ütemtervek megtalálása azonos specifikációjú számítógépes erőforrásokon (pl. Monte Carlo szimulációkra), amelyek adott feladatméret, fontosság és határidők mellett minimalizálják a teljes súlyozott késedelmességet (TWT).

Annak ellenére, hogy a felvetett két probléma látszólag eltérő tőről fakad, összefüggnek egymással a pénzügyi számítástudomány területén, továbbá a gyors, heurisztikus megközelítő megoldások megtalálása és alkalmazása NP nehéz problematikákra, közös. Eredményként téziseimmel újszerű és fontos hozzájárulásokat tettem a pénzügyi számítástudomány szakterületéhez, gyakorlati alkalmazások bemutatása révén.

1.1 Technológiai motiváció és elért eredmények

Tézisfüzetem e részében áttekintést adok a választott problémák motivációiról és gyakorlati alkalmazásairól, valamint tömören bemutatom elért kutatási eredményeimet.

1.1.1 Ritka, átlaghoz visszatérő portfólió kiválasztás

Amióta csak Markowitznak e jelentős fejlődést elindító munkája [32] megjelent az optimális kockázatszabályozott jövedelmezőséggel bíró portfóliók kiválasztásáról, azóta mind a kutató, mind a gyakorló csoportok körében aktív kutatási területnek számít. Ugyanakkor az átlaghoz visszatérés, mint az előrejelezhetőség klasszikus indikátora, ugyancsak rendkívüli figyelmet kapott az utóbbi évtizedek során. Kimutatták, hogy a hosszútávú tőzsdei többlethozamok az átlaghoz visszatérő jellegűek, éppen ezért tartalmazzák az előrejelezhetőség elemeit [21,31,36]. Miközben egyszerű és megbízható módszerek állnak rendelkezésre ahhoz, hogy egyelemű idősorokban visszatérési átlagot állapítsunk meg, addig a többdimenziós adatok alapján választott hasonló tulajdonságú portfólió kiválasztása sokkal nehezebb feladat. Ezt a Box-Tiao eljárás [15] segítségével tudjuk megoldani annak érdekében, hogy kointegrált vektorokat nyerjünk ki, általánosított sajátérték feladat megoldásával.

A közelmúltban publikált munkájában d'Aspremont [18] felvetette az átlaghoz visszatérő, ritkás kitöltésű portfóliók kérdéskörét. Ennek gyakorlati alkalmazása az a lehetőség, hogy olyan jövedelmező *konvergencia kereskedési stratégiát* alakítsunk ki, amelyekben akkor vásárolunk portfóliót, amikor ára a hosszútávú átlag alatt áll és akkor adjuk el, amikor ezen érték fölé emelkedett. A portfólió ritkasága ahhoz szükséges, hogy a konvergencia kereskedéssel járó tranzakciós költségeket csökkenthessük, ugyanakkor pedig emeljük a portfólió átláthatóságát. Szerzőnk új megközelítést alkalmazott a probléma megoldására a szemidefinit relaxáció bevezetésével és összehasonlította e megoldás hatékonyságát egy egyszerű mohó algoritmussal számos piacon. D'Aspremont feltételezi hogy a modell alapjául szolgáló folyamatok elsőrendű vektor autoregressziv folyamatot VAR(1) követnek és historikus adatokat használ a modell paramétereinek becsléséhez. A VAR(1) folyamatok parameter-becslésének jelentős irodalma ismert, a legújabb kutatások a ritkaságra és a szabályozott kovariancia becslésére összpontosítanak [10,17,38].

Az első téziscsoportban új, sűrű kitöltésű becslési módszereket vizsgáltam a megfigyelt VAR(1) folyamatok kovariancia mátrixának meghatározására. Vizsgálok olyan módszereket is, melyek az Ornstein-Uhlenbeck modell paramétereit határozzák meg. Végezetül új heurisztikus módszereket vizsgálok a számosság kényszere alatti optimális, leginkább előrelátható portfólió kiválasztására.

1.1.2 Optimális ütemezés azonos gépeken

Az ütemezési elmélet egyik konkrét alkalmazása a nagyméretű Monte Carlo szimulációk futtatása során merül fel, melyet pénzügyi szolgáltatásokban használnak komplex termékek árazása és kockázati mutatóinak kiszámitása érdekében. Eredményeként olyan valósidejű, optimális ütemezést találunk, amely minimalizálja a megbízások teljesítési idejét a kapacitás-beli megkötéseket figyelembe véve. Még specifikusabban, a pénzügyi számítástudomány területén a portfólió kiválasztás, a komplex pénzügyi eszközök árazása és kockázatkezelése elképesztő mértékű számítógépes erőforrást igényel, amelynek optimális használata alapvető fontosságú a befektetési pénzintézeteknél. A komplex portfóliók ár- és kockázatérzékenységi mutatóit naponta szükséges újraértékelni, erre ütemeznek nagyszámú éjszakai kalkulációt pénzügyi tranzakciók millióira, amihez számítógépek ezreit vetik be. Minden egyes megbízás jól megadott prioritással és végrehajtási határidővel ellátott, amelynek eredményeit a traderek, a kockázatkezelők és külső szabályozók is felhasználják. A megbízásokat általában le lehet állítani és újra lehet indítani egy későbbi időpontban, akár egy másik gépről is, amit a szakirodalom *leállítható-újraindítható tulajdonság*nak nevez. A probléma modellezési egyszerűsítése érdekében feltételezzük, hogy a gépek azonosak, és ismert, konstans mennyiségű gép áll rendelkezésre. Egyedi megbízási késedelmességet valamely megadott ütemezésben úgy határozunk meg, mint az az időmennyiség, amennyivel a megbízás később teljesül az előre meghatározott határidőnél, míg zérónak akkor tekintjük, ha megadott határidőre vagy már azt megelőzően teljesült.

Az elmúlt három évtized során behatóan tanulmányozták az azonos gépeken futó megbízások optimális ütemezését. Sahni [40] a *megvalósítható* ütemezés megtalálásához egy olyan $O(n \log mn)$ algoritmust mutat be, amely minden határidőt kielégít, már amennyiben van ilyen ütemezés, n megbízáshoz és m gépre. Az algoritmus alapgondolata szerint a megbízásokat azok legkorábbi esedékességi időpontja szerint ütemezi elsősorban, de a fennmaradó kapacítást kisebb megbízásokkal tölti ki. Megjegyezzük, hogy e módszer lehetővé teszi olyan algoritmus létrehozását, amely megtalálja a legminimálisabb gépi kapacítást, amelyhez egy megvalósítható ütemezés tartozik. Ezt az eredményt kiterjesztették azonos funkcionalitású, ám eltérő folyamatsebességű, úgynevezett *uniform gépekre* és olyan megbízásokra, amelyeknek kezdő- és határideje egyaránt adott; Martel [33] pedig polinom időben alkot meg erre a problémára egy elfogadható ütemezést, amennyiben van ilyen. Az ütemezési feladat ugyanakkor jóval nehezebbé válik, ha nincs megvalósítható ütemezés, de a célkitűzés az, hogy minimalizáljuk bizonyos mulasztások mértékét, amit gyakran *késedelmesség*ként definiálnak.

A valamennyi megbízásra érvényes maximális késedelmesség minimalizálására Lawler [27] bemutatja, hogy a kérdés megoldható polinom időben, még a feladatok sorrendjére vonatkozó megkötések mellett is. Martel [33] hasonló módon, polinom idejű algoritmust javasol olyan ütemezés létrehozására, amely minimalizálja a maximális késedelmességet. Sajátos, hogy amennyiben a mérték a teljes, és nem a maximális késedelmességre vonatkozik, abban az esetben a problémát egyetlen gépen is NP-nehéznek bizonyította Du és Leung [19]. Erre a problémára dinamikus programozás segítségével Lawler [26] kifejlesztett egy pszeudopolinom idejű algoritmust, ám ez csak az 1-gépes problémára vonatkozik és nem rendelkezik jó gyakorlati futásidő tulajdonsággal.

Amint a TWT probléma NP-nehézsége bizonyitva lett, a vonatkozó kutatások többsége gyors, heurisztikus algoritmusok kidolgozását vette célba. Dogramaci and Surkis [20] egyszerű heurisztikát javasol a teljes (nem-súlyozott) késedelmesség problémájára, leállítható-újraindítható tulajdonság nélkül. Rachamadugu and Morton [37] később azonos gépekre vizsgálta a teljes súlyozott késedelmesség problémáját, leállítható-újraindítható tulajdonság nélkül. Ők rövidlátó heurisztikát javasolnak, amelyet a legkorábbi határidős ütemezéshez (EDD), súlyozott legrövidebb folyamatidőhöz (WSPT) és Montagne szabályához hasonlítanak, kis probléma méret mellett (összesen 2 vagy 5 megbízás). Azizoglu és Kirca [12] algoritmust dolgoztak ki nem-súlyozott teljes késedelmességre leállítható-újraindítható tulajdonság nélkül, ám az ő algoritmusuk túlságosan lassú a gyakorlatban 15-nél több megbízás esetén. Armentano és Yamashita [11] nem-súlyozott problémát vizsgált leállítható-újraindítható tulajdonság nélkül, Koulamas [25] KPM heurisztikájából kiindulva és azt továbbfejlesztve, a tabu keresést felhasználva. Guinet [23] szimulált lehűtést alkalmaz a probléma megoldására azonos és uniform gépek esetében is és alsó korlátot ad az ajánlott módszer teljesítményének értékelésére. Még újabban Sen et al. [41] tekintette át a meglévő heurisztikus algoritmusokat az egy gépre vonatkozó teljes késedelmesség és teljes súlyozott késedelmesség problematikáját, míg Biskup et al. [14] ugyanezt végezte el azonos gépek teljes késedelmességi problematikájára és új heurisztikát is javasolt. Akyol és Bayhan [10] nagyszerű áttekintését adják a mesterséges neurális hálózat alapú megközelítéseknek az ütemezéssel összefüggésben és egy kapcsolt gradiens hálózatot javasolnak a súlyozott koraiság és késedelmesség összegének optimalizálására a több gépes esetben. A módszer megvalósíthatóságát egy 8 megbízásos ütemezéssel illusztrálják.

Magam egy újszerű heurisztikát terveztem a TWT problémára, melynek alapja a Hopfield neurális hálózati (HNN) megközelítés, mely jobban teljesít a meglévő egyszerű heurisztikáknál és kedvező skálázási tulajdonságokkal rendelkezik. Maheswaran et al. [28] hasonló megközelítést alkalmazott az egy gépes TWT problémára, eredményei pedig reményt keltőek egy konkrét 10 megbízásos problémára való alkalmazáskor.

Második téziscsoportomban újszerű, polinom idejű heurisztikákat vizsgálok az NP-nehéz TWT probléma megoldása érdekében. További lehetőségeket is vizsgálok a továbbfejlesztés érdekében véletlenszerű perturbációk felhasználásával. Végezetül tesztelem az újonnan kifejlesztett módszereket nagyszámú véletlenszerűen generált problémán, továbbá valós ütemezési feladatokon, amelyek a világ egyik vezető pénzügyi intézete, a Morgan Stanley jóvoltából állt rendelkezésemre.

2. A kutatásban alkalmazott modellek, módszerek

Annak érdekében, hogy a jelen disszertációban bemutatott eredményeket megkapjuk, számos modellt és számítási módszert használtunk fel és fejlesztettünk ki. Ebben a fejezetben ezeket mutatom be.

Ahhoz, hogy megoldhassuk a ritkás kitöltésű, optimálisan átlaghoz visszatérő portfóliók problematikáját a [18]-ban vázolt felépítést követem. D'Aspremont számos feltételt tett a model kifejlesztése során. Alapvető folyamatként fogadta el a VAR(1) modellt valamint azt, hogy az átlaghoz visszatérő portfóliókat az Ornstein-Uhlenbeck egyenlettel írhatjuk le. A modell paramétereinek meghatározása érdekében a Moore-Penrose pszeudoinvezet használtam az egyenlet megoldásához, amely a rekurziós mátrix maximális valószínűség megbecsüléséből következett, és kidolgoztam egy újszerű rekurzív módszert a kovariancia mátrix becslésére. Miután kiszámítottuk a VAR(1) folyamat paramétereit leképeztem az előrejelezhetőség maximalizálásának feladatát egy általánosított sajátérték problémává kardinalitási megkötésekkel, amely bizonyítottan NP nehéz [34]. Ezért szimulált lehűtést alkalmaztam annak érdekében, hogy polinóm idejű megközelítést találjak, amelynek hatékonyságát számos egyéb módszerrel vetettem össze: kimerítő és mohó kereséssel és egy újszerű heurisztikával, amelyet csonkítás elvén dolgoztam ki. Kiválasztva a portfóliót, újszerű módszert dolgoztam ki arra, hogy megtaláljam annak hosszútávú átlagát, amihez mintaillesztést használtam. Végezetül pedig a módszer gazdasági életképességét igazolandó, numerikus szimulációt futtattam újszerű konvergencia kereskedési stratégiát használva, amely döntéselméleti megfogalmazáson alapszik.

Az azonos gépeken futó optimális ütemezés, amely minimalizálja a teljes súlyozott késedelmességet, feladatának megoldása érdekében bináris ütemezési mátrix modellt alkalmaztam az ütemezés ábrázolásához. Tanulmányoztam továbbá más meglévő heurisztikus módszereket (EDD, WSPT, LWPF, LBS), amelyeket összehasonlítottam saját kifejlesztésű módszereimmel. Mátrixalgebrai transzformációkat használtam a probléma kvadratikus formává konvertálásához, a kényszereket beépítve a célfüggvénybe heurisztikus konstansok használatával. Ilyenformán lehetőségem nyílott a Ljapunov konvergencia alkalmazására a Hopfield Neurális Hálózaton keresztül annak érdekében, hogy polinom idejű közelítéses megoldásokat kapjak erre az NP nehéz problémára. A véletlenszerű perturbációk bevezetése révén és az ésszerűen kijelölt kezdőpontok segítségével még javítottam is ezen a módszeren.

Valamennyi módszer és modell implementálása MATLAB számítógépes környezetben történt, ahol numerikus szimulációk futottak mind mesterségesen előállított és valós, történeti adatokon, mindkét problémához. Disszertációmban bemutatott eredményeim alapját ezek a numerikus szimulációk adják.

felhasznált modellek	VAR(1) modell Ornstein-Uhlenbeck modell	Bináris ütemezési mátrix
alkalmazott módszerek	maximális valószínűség becslés Moore-Penrose pszeudoinverz újszerű rekurziós és kovariancia mátrix becslés benchmark módszerek (alapos és mohó kutatás, csonkításos módszer) szimulált lehűtés minta illesztésen alapuló, újszerű hosszútavú átlag becslés döntéselméleti megfogalmazáson alapuló konvergencia kereskedés	Heurisztikus optimalizálási módszerek (EDD, WSPT, LWPF, LBS) kvadratikus formák kombinatorikus optimalizálása Ljapunov konvergencia Hopfield Neurális Hálózat véletlenszerű perturbáció
ellenőrzések	numerikus szimulációk szintetikus és valós történeti adatokra MATLAB-on	
eredmények	1. téziscsoport	2. téziscsoport

1. ábra. Felhasznált modellek, módszerek és ellenőrzések, melyek e disszertáció eredményeihez elvezetnek.

3. Új tudományos eredmények

I. Téziscsoport: Új módszert fejlesztettem ki a VAR(1) folyamat paramétereinek becslésére, amely modellilleszkedési mértéket is ad. Mintailleszkedésen alapuló módszert vezetek be továbbá az Ornstein-Uhlenbeck folyamat hosszú távú középértékének becslésére. Ugyancsak új módszert alkalmazok a portfólió választásban a szimulált lehűtés adoptálása révén, végezetül bemutatok egy egyszerű konvergencia kereskedési algoritmust, így mutatván be kifejlesztett módszereim gyakorlati életképességét.

(E téziscsoporthoz kapcsolódó publikációim: 1, 2, 5, 6)

D'Aspremont konstrukcióját [18] követve a termékárat stacionárius, elsőrendű vektor autogresszív VAR(1) folyamatnak tekintem. Jelölje $s_{i,t}$ az i termék árát t időpillanatban, ahol $i = 1,...,n$ és t pozitiv egész számok s feltételezzük, hogy $\mathbf{s}_t^T = (s_{1,t},...,s_{n,t})$ egy VAR(1) folyamat, tehát:

$$\mathbf{s}_t = \mathbf{A}\mathbf{s}_{t-1} + \mathbf{W}_t, \tag{1}$$

ahol is **A** egy *nxn* mátrix és $\mathbf{W}_t \sim N(0,\sigma I)$ független, azonos eloszlású zaj, adott $\sigma > 0$. Jelöljük az $\mathbf{s}_t$ folyamat stacionárius kovariancia mátrixát **G**-vel feltételezvén, hogy az $\mathbf{s}_t$ történeti idősorozatok megfigyelhetők s a feladat az **A** és **G** mátrixok becslése.

Elsőként azt figyeltem meg, ha az n termékek száma nagyobb vagy egyenlő m-mel, a megfigyelt idősorok hosszával, akkor **A**-t az alábbi elsőfokú lineáris egyenletrendszer megoldása révén kapjuk:

$$\hat{\mathbf{A}}\mathbf{s}_{t-1} = \mathbf{s}_t. \tag{2}$$

Megjegyezzük, hogy ha $n > m$, ez a rendszer alulhatározott és végtelenül sok megoldása van. Ez esetben az első m terméket tartalmazó alrendszer egyértelmű megoldást ad. Tökéletes VAR(1) illeszkedést kapunk olyan idősorokban, ahol a potenciális termék portfólió nagyméretű a megfigyelt adatokhoz képest (pl. amely tartalmazza egyhavi időszak azon 500 papírjának napi

záróárát, amelyeket az S&P 500 index alatt jegyeznek), amiből egy ritkás kitöltésű, átlaghoz visszatérő portfólió választandó ki.
Viszont a legtöbb alkalmazásban az elérhető történeti idősorok hossza nagyobb a tekintetbe vett termékek számánál, vagyis (2) felülhatározott, és A-t például a minimum négyzetes hibabecslési módszer alkalmazásával becsülhetjük, mint alább

$$\hat{\mathbf{A}} = \arg\min_{\mathbf{A}} \sum_{t=2}^{m} \left\| \mathbf{s}_t - \mathbf{A}\mathbf{s}_{t-1} \right\|^2 \tag{3}$$

ahol $\|\cdot\|$ az euklideszi normát jelenti.

A fenti kifejezés az **A** mátrix minden egyes elemére vett parciális deriváltjainak gyökeinek meghatározására az alábbi egyenletrendszert kapjuk:

$$\sum_{k=1}^{n} \hat{\mathbf{A}}_{i,k} \sum_{t=2}^{m} s_{k,t-1} s_{j,t-1} = \sum_{t=2}^{m} s_{i,t} s_{j,t-1} \forall i, j = 1,\dots,n. \tag{4}$$

$\hat{\mathbf{A}}$ -ra megoldva és visszatérve **s** vektorjelöléséhez az alábbit kaptam

$$\hat{\mathbf{A}} = \sum_{t=2}^{m} \mathbf{s}_t \mathbf{s}_{t-1}^T \left(\sum_{t=2}^{m} \mathbf{s}_{t-1} \mathbf{s}_{t-1}^T \right)^+, \tag{5}$$

ahol $\mathbf{M}^+$ az **M** mátrix Moore-Penrose pszeudoinverze. Megjegyezzük, hogy a Moore-Penrose pszeudoinverzet alkalmaztam, hogy az $\mathbf{s}_{t-1}\mathbf{s}_{t-1}^T$ esetlegesen felmerülő szingularitási problémáját kikerüljem.

Feltételezve, hogy az (1) egyenletben a zaj tagok független, azonos eloszlásúak $\mathbf{W}_t \sim N(0, \sigma I)$ adott $\sigma > 0$ -ra, az alábbi becsléshez jutottam $\hat{\mathbf{A}}$ -t használva az (5)-ből:

$$\hat{\sigma} = \sqrt{\frac{1}{n(m-1)} \sum_{t=2}^{m} \left\| \mathbf{s}_t - \hat{\mathbf{A}}\mathbf{s}_{t-1} \right\|^2}. \tag{6}$$

Abban az általánosabb esetben, ha $\mathbf{W}_t$ zaj tagjai korreláltak, a zaj **K** kovariancia mátrixát az alábbiak szerint becsülöm meg:

$$\hat{\mathbf{K}} = \frac{1}{m-1} \sum_{t=2}^{m} (\mathbf{s}_t - \hat{\mathbf{A}}\mathbf{s}_{t-1})\ (\mathbf{s}_t - \hat{\mathbf{A}}\mathbf{s}_{t-1})^T. \tag{7}$$

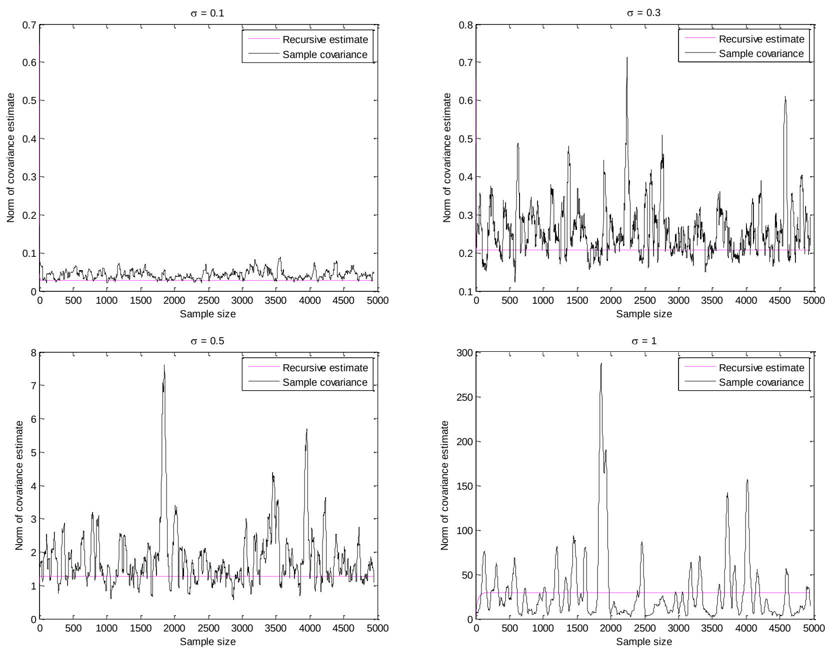
A megbecsült zaj kovariancia mátrixát később az egész folyamat kovariancia mátrixának megbecsüléséhez használom.

I.1. Új numerikus módszert dolgoztam ki a historikus idősorok kovariancia mátrixának becslésére, VAR(1) modellt feltételezve, amely gyors és megbízható becslést ad, egy modellilleszkedési mértékkel egyetemben.

A módszer az alábbi képleten alapul:

$$\mathbf{G}(k+1)=\mathbf{G}(k)-\delta(\mathbf{G}(k)-\mathbf{A}\mathbf{G}(k)\mathbf{A}^{T}-\mathbf{K}), \tag{8}$$

ahol δ egy konstans 0 és 1 között, $\mathbf{G}(k)$ pedig a k-ik iterációban becsült kovariancia mátrix. Biztosítván, hogy a numerikus módszer kezdőpontja, $\mathbf{G}(0)$ pozitív definit (pl. a minta kovariancia mátrix) s mivel a becsült $\mathbf{K}$ pozitiv definit, ez az iterative módszer pozitiv definit tulajdonságú becslést eredményez. Ez egyenletesebb konvergenciát ad a valós kovariancia mátrixhoz, mint a minta kovariancia mátrix, ahogyan azt az alábbi, generált adatok segítségével készített ábrán bemutatom.



2. ábra: Minta kovariancia és rekurzív kovariancia becslés 5000 mintán, 50-es méretű csúszóablakban σ =0.1, 0.3, 0.5, 1-re (az ábrák különböző skálázásúak)

E módszer egy modellilleszkedési mértéket is ad a fenti rekurzív becslést a minta kovariancia mátrixához hasonlítva. A konvergencia kereskedésben ez alkalmazható bizonyosság mértékként, hogy valóban egy átlaghoz visszatérő portfóliót találtunk-e.

A következő eredmény az Ornstein-Uhlenbeck egyenlet paramétereinek becslése, az alábbiak szerint

$$dp_t = \lambda\left(\mu - p_t\right)dt + \sigma dW_t, \tag{9}$$

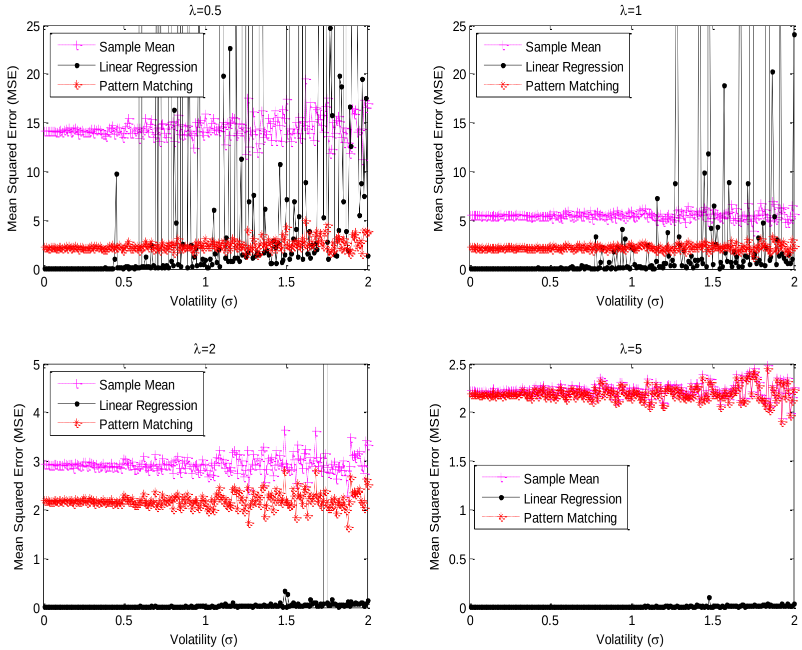
ahol W_t egy Wiener folyamat és $\lambda > 0$ (átlaghoz visszatérés paramétere), μ (hosszú távú átlag), és $\sigma > 0$ (portfólió volatilitás) konstansak.

I.2. Új módszert alkalmaztam az Ornstein-Uhlenbeck folyamat átlag becslésére, mintaillesztéses megközelítéssel. Ennek során a legvalószínűbb átlagot választjuk ki az alapjául szolgáló OU folyamatok gauszi sűrűségeinek maximalizálásával. Az új becslést az alábbi kifejezéssel kapjuk meg

$$\hat{\mu}_3 := \frac{\sum_{i=1}^{t}\sum_{j=1}^{t}\left(\mathbf{U}^{-1}\right)_{i,j}\left[\mu(0)\ \left(2e^{-\lambda(i+j)} - e^{-\lambda i} - e^{-\lambda j}\right) - 2\mathbf{p}_j\left(e^{-\lambda i} - 1\right)\right]}{\sum_{i=1}^{t}\sum_{j=1}^{t} 2\left(\mathbf{U}^{-1}\right)_{i,j}\left(e^{-\lambda(i+j)} - e^{-\lambda i} - e^{-\lambda j} + 1\right)}, \tag{10}$$

ahol $\mathbf{U}_{ij} := E\left(p(t-i)\,p(t-j)\right) = \frac{\sigma^2}{2\lambda}e^{-\lambda(i+j)}\left(e^{2\lambda i} - 1\right)$ $\mathbf{p}_t$ időkorrelációs mátrixa.

Ezzel a módszerrel egy a lineáris regressziónál stabilabb és a mintaátlagnál pontosabb becslést kapunk, ahogy azt alábbi ábránkon bemutatjuk.



3. ábra: MSE különböző μ becslésekre σ függvényeként. λ =0.5, 1, 2, 5 , $\mu_0 = 0$ és a minta mérete 20.

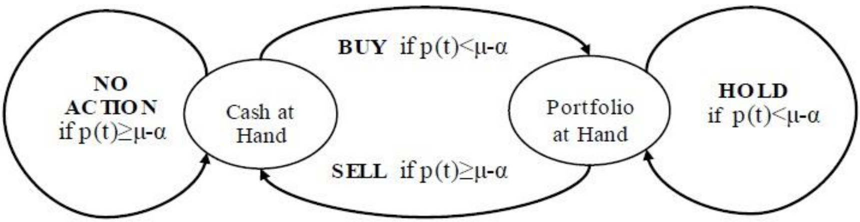
Az átlaghoz visszatérő portfólió meghatározása és a hosszú távú átlag μ megbecsülését követő fő feladat annak ellenőrzése vajon $\mu(t) < \mu$ vagy $\mu(t) \geq \mu$, a minták $\left\{ p(t) = \mathbf{x}^T \mathbf{s}(t), t = 1, \ldots, T \right\}$ alapján. Az ellenőrzést döntéselméleti problémaként foghatjuk fel, lévén hogy $\mu(t)$ közvetlen megfigyelései nem elérhetők.

Amennyiben a $p(t)$ folyamat stacionér állapotú, úgy a minták $\left\{ p(t), t = 1, \ldots, T \right\}$ normális eloszlásúak: $N\left(\mu, \sqrt{\frac{\sigma^2}{2\lambda}} \right)$.Tehát bármely megadott ε elfogadható hibahatárra, választhatok α –t, hogy

$$\int_{\mu-\alpha}^{\mu+\alpha} \frac{1}{\sqrt{2\pi\sigma^2/2\lambda}} e^{-\frac{(u-\mu)^2}{\sigma^2/\lambda}} du = 1-\varepsilon. \qquad (11)$$

I.3. Egyszerű kereskedési stratégiát dolgoztam ki mely döntéselméleti megfogalmazáson alapszik. Ebből egy konvergencia kereskedési algoritmust fejlesztettem ki, amely különböző portfóliókiválasztási algoritmusok összevetésére alkalmazható.

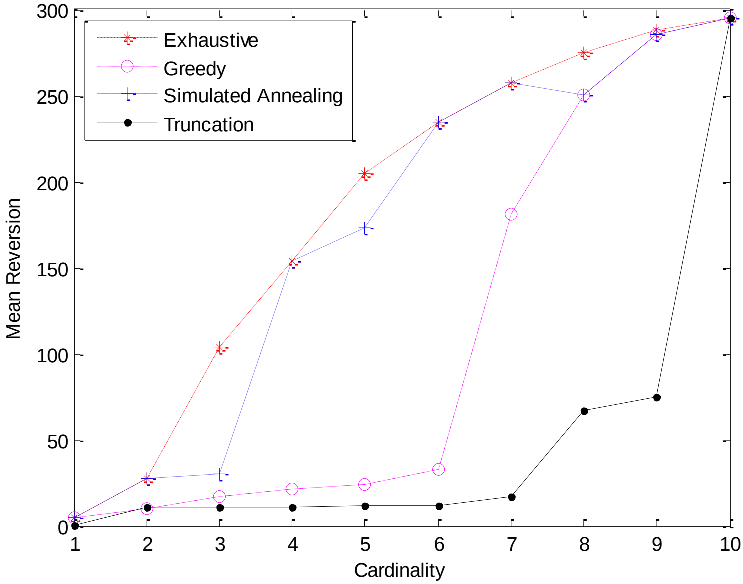
Az algoritmus minden időintervallumban ismétlődő lépései az alábbi ábrában kerültek összefoglalásra:



4. ábra: Átlaghoz visszatérő portfólió egyszerű konvergencia kereskedési folyamatábrája

I.4. Adoptáltam a szimulált lehűtés optimalizálási módszert az átlaghoz visszatérés maximalizálása érdekében ritkássági megkötés mellett. Konvergencia-szabályokat és szomszédkiválasztási algoritmust fejlesztettem ki, amelyek jól teljesítenek véletlenszerűen generált példákon. Ez a módszer konzisztensen felülmúlta a mohó heurisztikát a véletlenszerűen generált esetek 20%-ában, elfogadható futásidő mellett.

Bizonyos generált esetekben, a szimulált lehűtés felülmúlja a mohó módszert konzisztensen magasabb átlaghoz visszatérési paraméterrel a kardinalitás-beli megkötések mellett. Példát erre az alábbi ábrán adunk.



5. ábra: Különböző méretű portfóliókat kiválasztó módszerek összehasonlítása n=10-dimenziós generált VAR(1) adatokon.

Az elképzelés ellenőrzéséhez end-to-end implementáltam a paraméter becslési módszereket, a szimulációs lehűtésen alapuló portfólió kiválasztási algoritmust valamint a döntéselméleti kereskedési stratégiát. Azt találtam, hogy valós historikus adatokon, ezen új módszereket egymás után, összehangoltan alkalmazva, jelentős mértékű hozamtöbblet keletkezik a klasszikus módszerekkel szemben.

II. Téziscsoport: Bemutattam, hogy a TWT probléma hogyan konvertálható kvadratikus formává és hogyan írható le olyan speciális kvadratikus célfüggvény, melynek segítségével használhatjuk a Hopfield neurális hálózatot az TWT probléma megoldására. Bemutatom, hogyan teljesíti túl ez a módszer az egyszerűbb heurisztikákat véletlenszerűen generált, nagyszámú problémán, majd intelligens kezdőpontok bevezetésével illetve véletlenszerű perturbációk hozzárendelésével tovább javítom a módszert. Bemutatom e módszerek gyakorlati alkalmazhatóságát egy konkrét éjszakai ütemezési feladaton, mely a világ egyik legnagyobb pénzügyi intézeténél, a Morgan Stanleynél merült fel..

(E téziscsoporthoz kapcsolodó publikációim: 3, 4)

A teljes súlyozott késések (TWT) optimalizálásával az alábbi formalizmusban dolgoztam. Adott N feladat melyek mérete $\mathbf{x} = \{x_1, x_2, x_3, \ldots x_N\} \in \mathbb{N}^N$, a feladatok végrehajtását bármely időpillanatban leállíthatjuk és újraindíthatjuk, így az egyes feladataok végrehajtásának időegységei nem feltétlenül folyamatosak. A szakirodalom ezt a feltételt *leállítható-újraindítható tulajdonság*ként ismeri és feltételezi hogy valamely gépen elindított feladatot bármely másikon folytatni lehet [16]. Minden feladathoz előírt egy határidő, amely így írható le: $\mathbf{K} = \{K_1, K_2, K_3, \ldots, K_N\} \in \mathbb{N}^N$. E megkötés határozza meg, hogy mely feladatot meddig kell elvégezni. A processzorok konstans számát, a rendszer kapacitását V jelöli. Adott egy súlyvektor is $\mathbf{w} = \{w_1, w_2, w_3, \ldots, w_N\} \in \mathbb{R}^N$, $w_i \geq 0, \forall i = 1,\ldots, N$, mely a relatív prioritást (vagy a súlyt) jelöli valamennyi feladathoz, amelyet az célfüggvény definiálásához használunk. Az ütemezést ez a bináris mátrix jelöli: $\mathbf{C} \in \{0,1\}^{N \times L}$, ahol $C_{i,j} = 1$ ha az i feladat j időpillanatban feldolgozásra került, és L jelöli az ütemezés maximális hosszát. Példát is bemutatunk optimális ütemezésre (12), melynek paraméterei a következők: $V = 2$, $N = 3$, $\mathbf{x} = \{2,3,1\}$, $\mathbf{K} = \{3,3,3\}$.

$$\mathbf{C} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \tag{12}$$

Ahhoz hogy az adott **C** ütemezés hatékonyságát értékeljük, egy feladat *késedelmességét* az alábbiak szerint definiáljuk:

$$T_i = \max\left(0, F_i - K_i\right), \tag{13}$$

ahol F_i az i feladat aktuális befejezése a **C** ütemezésben: $F_i = \arg\max_j \left\{C_{i,j} = 1\right\}$ (Az utolsó 1 pozíciója az i-dik sorban a $\mathbf{C}$ ütemező mátrixban.)

A megoldásra váró probléma formálisan alábbiak szerint írható fel:

$$\mathbf{C}_{opt} := \arg\min_{\mathbf{C}} \sum_{i=1}^{N} w_i T_i \tag{14}$$

A következő megkötésekkel:

- A feladatok hossza az ütemezési mátrixban megegyezik a megadott összeggel

$$\sum_{j=1}^{L} C_{i,j} = x_i, \forall i = 1,\ldots,N \tag{15}$$

- Az ütemezett feladatok száma adott időpillanatban nem haladja meg a rendszer kapacitását:

$$\sum_{i=1}^{N} C_{i.j} \le V, \forall j = 1,\ldots,L \tag{16}$$

II.1. Igazoltam, hogy a TWT probléma kvadratikus optimalizációs problémává konvertálható, melyben benne foglaltatnak a megkötések is, heurisztikus együtthatókkal.

Az optimalizációs feladatot kvadratikus formába transzformáltam:

$$f\left(\mathbf{y}\right) = -\frac{1}{2}\mathbf{y}^T\mathbf{W}\mathbf{y} + \mathbf{b}^T\mathbf{y}, \tag{17}$$

ahol

$$\mathbf{W} = \alpha\mathbf{W}_A + \beta\mathbf{W}_B + \gamma\mathbf{W}_C \in \mathbb{R}^{NL\times NL} \tag{18}$$

és

$$\mathbf{b} = \alpha\mathbf{b}_A + \beta\mathbf{b}_B + \gamma\mathbf{b}_C \in \mathbb{R}^{NL\times 1}, \tag{19}$$

ahol A alsóindex megfelel a célfüggvénynek, a B és C alsóindexek a két megkötésnek és α, β, γ heurisztikus együtthatók.

A minimalizációs probléma megfelel:

$$\mathbf{C}_{opt} := \arg\min_{\mathbf{C}} \sum_{i=1}^{N} \sum_{j=K_i+1}^{L} w_i C^2{}_{i,j} \tag{20}$$

A leképezést alábbi egyenlet megoldása adja:

$$-\frac{1}{2}\mathbf{y}^T\mathbf{W}_A\mathbf{y} + \mathbf{b}_A{}^T\mathbf{y} = \min \sum_{i=1}^{N} \sum_{j=K_i+1}^{L} w_i C^2{}_{i,j} \tag{21}$$

A megoldás az alábbi:

$$\mathbf{b}_A = \mathbf{0}_{JL\times 1}, \tag{22}$$

$$\mathbf{W}_A = -2\begin{pmatrix} \mathbf{D}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{D}_J \end{pmatrix}, \tag{23}$$

ahol

$$\mathbf{D}_j = \begin{pmatrix} \mathbf{0}_{K_j\times K_j} & \mathbf{0}_{K_j\times L-K_j} \\ \mathbf{0}_{L-K_j\times K_j} & w_j * \mathbf{I}_{L-K_j\times L-K_j} \end{pmatrix} \in \mathbb{R}^{L\times L}. \tag{24}$$

Az első megkötés így fejezhető ki

$$-\frac{1}{2}\mathbf{y}^T\mathbf{W}_B\mathbf{y}+\mathbf{b}_B{}^T\mathbf{y}=\min\sum_{i=1}^{N}\left(\left(\sum_{j=1}^{L}C_{i,j}\right)-x_i\right)^2, \tag{25}$$

ahol a megoldás

$$\mathbf{b}_B=2\begin{pmatrix}x_{1_{1\times L}} & x_{2_{1\times L}} & \cdots & x_{J_{1\times L}}\end{pmatrix} \tag{26}$$

$$\mathbf{W}_B=-2\begin{pmatrix}\mathbf{1}_{L\times L} & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \mathbf{1}_{L\times L} & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{1}_{L\times L}\end{pmatrix}. \tag{27}$$

A második megkötés így fejezhető ki

$$-\frac{1}{2}\mathbf{y}^T\mathbf{W}_C\mathbf{y}+\mathbf{b}_C{}^T\mathbf{y}=\min\sum_{i=1}^{M}\left(\left(\sum_{j=1}^{N}C_{i,j}\right)-V\right)^2, \tag{28}$$

megoldása

$$\mathbf{b}_C=\left[\mathbf{V}_{M\times 1},\mathbf{0}_{L-M\times 1},\mathbf{V}_{M\times 1},\mathbf{0}_{L-M\times 1},\ldots,\mathbf{V}_{M\times 1},\mathbf{0}_{L-M\times 1}\right], \tag{29}$$

$$\mathbf{W}_C=-2\begin{pmatrix}\mathbf{D} & \mathbf{D} & \cdots & \mathbf{D}\\ \mathbf{D} & \mathbf{D} & \cdots & \mathbf{D}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{D} & \mathbf{D} & \cdots & \mathbf{D}\end{pmatrix}, \tag{30}$$

ahol

$$\mathbf{D}=\begin{pmatrix}\mathbf{I}_{M\times M} & \mathbf{0}_{M\times L-M}\\ \mathbf{0}_{L-M\times M} & \mathbf{0}_{L-M\times L-M}\end{pmatrix}. \tag{31}$$

II.2. A HNN módszert alkalmazva közelítő megoldást találtam a TWT ütemezési problémára, polinom idő alatt. Demonstráltam, hogy a HNN megoldás felülteljesíti az egyszerűbb heurisztikákat, amilyen az EDD, WSPT, LBS, random és LWPF, nagyszámú, véletlenszerűen generált TWT feladaton.

Előszeretettel alkalmazott, hatékony heurisztikus algoritmus a kvadratikus optimalizálási problémára a Hopfield Neural Network (HNN). E neurális hálózatot az alábbi iteratív átmeneti szabály írja le:

$$\mathbf{y}_i(k+1) = \operatorname{sgn}\left(\sum_{j=1}^{N} \hat{W}_{ij} y_j(k) - \hat{b}_i\right), i = \operatorname{mod}_N k, \tag{32}$$

ahol

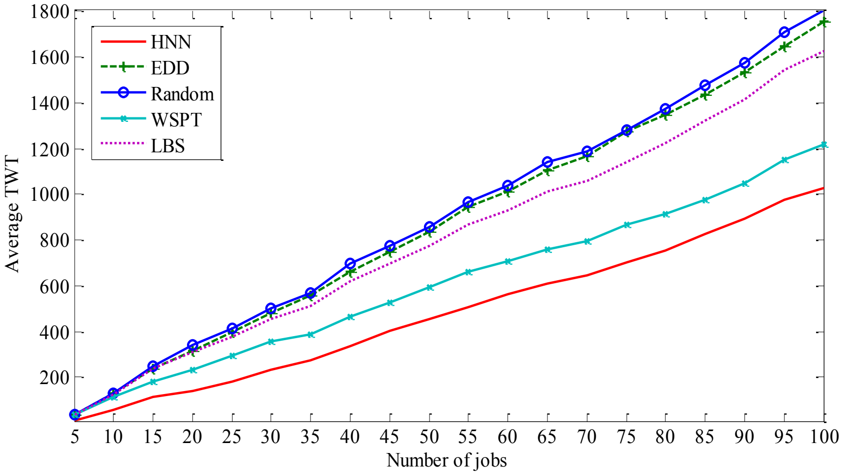
$$\begin{aligned} \mathbf{d} &= -\operatorname{diag}(\mathbf{W}) \\ \mathbf{W} &= -\mathbf{W} - \operatorname{diag}(\mathbf{d}) \\ \mathbf{b} &= \mathbf{b} - \frac{1}{2}\mathbf{d}. \end{aligned} \tag{33}$$

Ljapunov módszert alkalmazva, Hopfield [24] bizonyította, hogy a HNN konvergál annak fix pontjához, következésképp minimalizálja a kvadratikus Ljapunov függvényt:

$$\mathcal{L}(\mathbf{y}) := -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \hat{W}_{ij} y_i y_j + \sum_{i=1}^{N} y_i \hat{b}_i = -\frac{1}{2}\mathbf{y}^T \hat{\mathbf{W}} \mathbf{y} + \hat{\mathbf{b}}^T \mathbf{y}. \tag{34}$$

Tehát a HNN képes megoldani a kombinatorikus optimalizációs problémát polinom időben, speciális feltételek teljesülése esetén, ahogy azt [29] és [30] bemutatta.

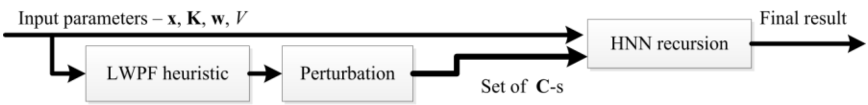
Teljeskörű implementációt végeztem MATLAB szimulációs környezetben és teszteltem a módszer hatékonyságát a szakirodalomban megtalálható módszerekkel szemben (EDD, SWPT, LBS, LWPF) nagyszámú, véletlenszerűen generált, különböző méretű problémára. Az eredményt az alábbi ábra mutatja:



6. ábra: Különböző heurisztikák TWT átlaga véletlenszerűen generált problémákon, a probléma méretének függvényeként.

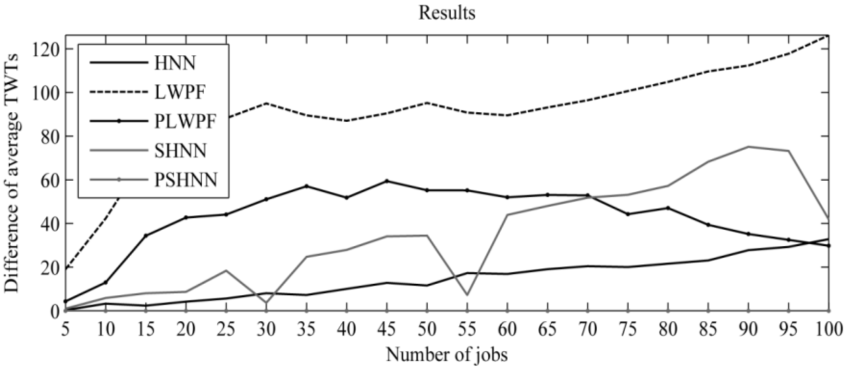
II.3. Továbbfejlesztettem a HNN módszert a TWT problémára a kezdőpontok intelligens megválasztása révén (SHNN) illetve e kezdőpontok véletlenszerű perturbációjával (PSHNN).

A perturbált, finomított Hopfield neurális hálózat (PSHNN) tömbdiagramja az alábbi:



7. ábra: A PSHNN módszer tömbdiagramja.

Azt találtam, hogy a PSHNN módszer konzisztensen felülteljesíti a többi módszert nagyszámú véletlenszerűen generált TWT feladaton, ahogy azt alábbi ábra mutatja.



8. ábra: Különböző heurisztikák által előállított átlag TWT és a legjobb módszer (PSHNN) TWT átlagának eltérései, 100 véletlenszerűen generált problémán, a probléma mértének függvényében.

Az elképzelés igazolására, alkalmaztam e módszert egy konkrét, valós, nagy méretű ütemezési problémára. Ebben több, mint 500 számítógépes feladat szerepel a Morgan Stanley-nél, a világ egyik legnagyobb pénzintézeténél, éjszaka futó, derivatív portfóliók kockázati érzékenységét számoló algoritmusok kapcsán. Eredményeinket az alábbi táblázat foglalja össze: a PSHNN 5 %-kal felülteljesíti a következő leghatékonyabb módszert a PLWPF-et.

Súly	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**ÖSSZEG**	**Relatív különbség PSHNN-hez képest**
PSHNN	4401	11116	4020	1620	1092	8	0	0	22 257	0%
PLWPF	3513	9624	5130	1788	490	312	2304	190	23 351	5%
HNN	4404	11040	4735	1824	1092	456	468	0	24 019	7%
LWPF	4404	11140	5470	2472	1183	40	0	0	24 709	10%
EDD	4401	9940	1770	636	1134	464	22752	1430	42 527	48%

9.ábra: A különböző súlyú feladatok összes súlyozott késését összefoglaló táblázat a különböző heurisztikákra, a Morgan Stanley-nél jelentkező ütemezési feladatra.

4. Következtetések és az eredmények alkalmazásai

Jelen disszertációban a pénzügyi számítástechnika két fontos szakterületét vizsgáltam, a portfólió kiválasztást és az optimális erőforrás ütemezést. NP nehéz nyílt problémát választottam mindkét szakterületről és polinom idejű heurisztikus módszereket vizsgáltam olyan gyors megoldások elérése érdekében, melyek túlteljesítenek egyéb hasonló módszereket. Minden esetben igyekeztem új megközelítéseket bevezetni, amelyek

- jobb értékeket kínálnak más, a szakirodalomban fellelhető heurisztikus módszerekhez képest
- alacsony számítási komplexitásúak, tehát polinom futási idejűek
- megfelelő futásidő tulajdonsággal rendelkeznek, így a gyakorlati feladatvégrehajtásban is alkalmazhatók.

Sikerült továbbá számos egyéb hozzájárulást tennem minden felvetett probléma megoldásához. A ritkás kitöltésű, átlaghoz visszatérő portfólió kiválasztás kérdésében jelentős előrelépést sikerült elérnem a VAR(1) és az Ornstein-Uhlenbeck modell paramétereinek becslésében. A TWT ütemezés problémakörében igazoltam, hogy az konvertálható kvadratikus optimalizációs problémává nem triviális mátrix algebrai leképezéssel. A klasszikus HNN módszer alkalmazásán túl sikerült előrelépnem véletlenszerű perturbációk bevezetésével, intelligensen megválasztott kezdőpontokra (PSHNN módszer).

Tekintettel e fenti eredményekre, disszertációmban kitűzött céljaimat sikerült elérnem.

Nem utolsó sorban, minden esetben elvégeztem az elképzelés ellenőrzését, és széleskörű szimulációkat is végeztem generált és valós adatokon egyaránt. A valós adatok eredményei minden esetben meggyőzőek, hiszen 34%-os kereskedési nyereség demonstrálható a historikus S&P 500 pénzpiacon átlaghoz visszatérő portfóliók konvergens kereskedését alkalmazva, egyúttal 10%-os előrelépést tudtam bemutatni a legjobbnak mondható, a szakirodalomban bemutatott heurisztikához képest, valós és hosszú ütemezésű TWT probléma megoldásában, amely a Morgan Stanley-nél jelentkezett.

5. Összefoglalás

A disszertáció új eredményekkel járul hozzá a pénzügyi számítástudomány fejlődéséhez új algoritmusok bevezetése révén a kombinatorikus optimálási problematikák egy osztályára, melybe beletartoznak a ritkás kitöltésű portfóliók optimalizálása (az előreláthatóság maximálása kardinalitás-beli megkötések mellett) továbbá az ütemezési optimalizálás (a teljes súlyozott késedelmesség minimalizálása azonos gépeken, teljesítmény-megkötöttségek mellett).

A valós problematikák numerikus teszteredményeit az alábbi táblázat foglalja össze:

terület	*valós probléma*	*átlag teljesítés hagyományos megközelítéssel*	*átlag teljesítés a javasolt új módszerrel*	*hatása a pénzügyi számítástud-ra (százalékos javulásban)*
portfólió optimalizálás	konvergens kereskedés az USA S&P 500 pénzpiacon	11.6% (éves S&P 500 index hozam)	34%	22.4%
ütemezés optimalizálás	Morgan Stanley éjszakai ütemezési probléma	24709 (LWPF teljesítmény)	22257 (PSHNN teljesítmény)	10%

6. Köszönetnyilvánítás

Tisztelettel kívánom ezennel megköszönni témavezetőm folyamatos iránymutatását a kutatómunkám során. Köszönöm továbbá Tornai Kálmán, Sipos Róbert, Farhad Kia, dr. Jeney Gábor és dr. Iványi Antal együttműködését és a velük folytatott rendkívül hasznos eszmecseréket, amelyeket a tézisekben bemutatott tárgykörök kapcsán folytathattunk. Köszönöm a Morgan Stanleynek, munkaadómnak, aki számos távollétet biztosított részemre e munka elkészítéséhez, illetve adatokat bocsátott rendelkezésemre a módszerek hatékonyságának vizsgálatához, valós alkalmazásokban. Végezetül, köszönöm feleségemnek, Ildikónak, s egész családomnak a munka során mindvégig nyújtott támogatást.

7. Szerző publikációs jegyzéke

Folyóirat publikációk [19 pont]

1. Fogarasi, N., Levendovszky, J. (2012) A simplified approach to parameter estimation and selection of sparse, mean reverting portfóliós. *Periodica Polytechnica,* 56/1, 21-28. [4 pont]

2. Fogarasi, N., Levendovszky, J. (2012) Improved parameter estimation and simple trading algorithm for sparse, mean-reverting portfóliós. *Annales Univ. Sci. Budapest., Sect. Comp.*, 37, 121-144. [4 pont]

3. Fogarasi, N., Tornai, K., & Levendovszky, J. (2012) A novel Hopfield neural network approach for minimizing total weighted tardiness of jobs scheduled on identical machines. *Acta Univ. Sapientiae, Informatica*, 4/1, 48-66. [6/2=3 pont]

4. Tornai, K., Fogarasi, N., & Levendovszky, J. (2013) Improvements to the Hopfield neural network solution to the total weighted tardiness scheduling problem. *Periodica Polytechnica,* 57/2, 57-64. [4/2=2 pont]

5. Fogarasi, N., Levendovszky, J. (2013) Sparse, mean reverting portfólió selection using simulated annealing. *Algorithmic Finance,* 2/3-4, 197-211. [6 pont]

Konferencia előadások

6. Fogarasi, N., Levendovszky, J. (2012) Combinatorial methods for solving the generalized eigenvalue problem with cardinality constraint for mean reverting trading. *9th Joint Conf. on Math and Comp. Sci. February 2012 Siófok, Hungary*

Technikai beszámolók

7. Fogarasi, N. (2003) Option pricing primer. *Technical Report*
8. Fogarasi, N. (2003) Non-gaussian option pricing. *Technical Report*
9. Fogarasi, N. (2003) Option pricing using neural networks. *Technical Report*

8. A disszertációban hivatkozott publikációk

10. Akyol, D.E., & Bayhan, G.M. (2008) Multi-machine earliness and tardiness scheduling problem: an interconnected neural network approach. *International Journal of Advanced Manufacturing Technology*, Springer-Verlag London Limited 37:576-588
11. Armentano, V.A., & Yamashita, D.S. (2000) Taboo search for scheduling on identical parallel machines to minimize mean tardiness. *Journal of Intelligent Manufacturing*, 11, 453-460.
12. Azizoglu, M., & Kirca, O. (1998) Tardiness minimization on parallel machines. *International Journal of Production Economics,* 55, 163-168.
13. Banerjee, O., El Ghaoui, L., & d'Aspermont A. (2008) Model Selection Through Sparse Maximum Likelihood Estimation. *Journal of Machine Learning Research,* 9: 485-516
14. Biskup, D., Herrman, J., & Gupta, J.N.D. (2008) Scheduling identical parallel machines to minimize total tardiness. *International Journal of Production Economics,* 115, 134-142
15. Box, G.E., & Tiao, G.C. (1977) A canonical analysis of multiple time series. *Biometrika* 64(2), 355-365
16. Brucker, P. (2007) *Scheduling Algorithms*. Fifth Edition, New York: Springer
17. D'Aspermont, A, Banerjee, O., & El Ghaoui, L.(2008) First-Order Methods for Sparce Covariance Selection. *SIAM Journal on Mátrix Analysis and its Applications,* 30(1) 56-66
18. D'Aspremont, A.(2011) Identifying small mean-reverting portfóliós. *Quantitative Finance,* 11:3, 351-364
19. Du, J. & Leung, J. Y.-T. (1990) Minimizing total tardiness on one machine is NP-hard. *Math. Oper. Res*., 15(3), 483-495.
20. Dogramaci, A., & Surkis, J. (1979) Evaluation of a heuristic for scheduling independent jobs on parallel identical processors. *Management Science,* 25(12), 1208-1216.
21. Fama, E., & French, K.(1988) Permanent and Temporary Components of Stock Prices. *The Journal of Political Economy* 96(2), 246-273
22. Grant, Jeremy (2010) High-frequency trading: Up against a bandsaw. *Financial Times* 2010 Sep 2.
23. Guinet, A. (1995) Scheduling independent jobs on uniform parallel machines to minimize tardiness criteria. *Journal of Intelligent Manufacturing*, 6, 95-103.
24. Hopfield, J. J. (1982) Neural networks and physical systems with emergent collective computational abilities in *Proceedings of the National Academy of Sciences of the USA*

25. Koulamas, C.P. (1994) The total tardiness problem: review and extensions. *Operations Research*, 42, 1025-1041.
26. Lawler, E. L. (1977) A pseudopolynomial algorithm for sequencing jobs to minimize total tardiness. *Annals of Discrete Mathematics*., 1, 331-342
27. Lawler, E. L. (1981) Preemptive scheduling of precedence-constrained jobs on parallel machines, deterministic and stochastic scheduling. *Proceedings of the NATO Advanced Study and Research Institute on Theoretical Approaches to Scheduling Problems*
28. Maheswaran, R., Ponnambalam, S.G., Nithin S. D., & Ramkumar, A.S. (2004) Hopfield Neural Network Approach for Single Machine Scheduling Problem. *Proceedings of the 2004 IEEE Conference on Cybernetics and Intelligent Systems*
29. Mandziuk, J. (1996) Solving the Travelling Salesman Problem with a Hopfield - type neural network. *Demonstratio Mathematica*, Vol. 29(1)
30. Mandziuk, J.,& Macukow, B. (1992) A neural network designed to solve the N-Queens Problem. *Biological Cybernetics*, Vol 66(4)
31. Manzan, S. (2007) Nonlinear Mean Reversion in Stock Prices. *Quantitative and Qualitative Analysis in Social Sciences*, 1(3), 1-20
32. Markowitz, H. (1952) Portfólió Selection. *The Journal of Finance* 7:1,77-91
33. Martel, C. (1982) Preemptive Scheduling with Releast Times. *Deadlines and Due Dates Journal of the Association for Computing Machinery*. Vol 29, No 3, 812-829
34. Natarjan, B.K. (1995) Sparse approximate solutions to linear systems. *SIAM J. Comput.* 24(2), 227-234
35. Pinedo, M. (2008) *Scheduling – Theory, Algorithms and Systems*. Third Edition, New York: Springer
36. Poterba, J.M., & Summers, L.H. (1988) Mean reversion in stock prices: Evidence and implications. *Journal of Financial Economics* 22(1), 27-59
37. Rachamadugu, R.M.V., & Morton, T.E. (1983) Myopic heuristics for the weighted tardiness problem on identical parallel machines.*Tech Report CMU-RI-TR-83-17*, Carnegie-Melon University, the Robotics Institute.
38. Rothman, A., Bickel, P., Levina, E., & Zhu, J. (2008). Sparse permutation invariant covariance estimation. *Electronic Journal of Statistics*, 2: 494-515
39. Sahni, S. (1974) Computationally related problems. *SIAM Journal on Computing*, 3.
40. Sahni S. (1979) Preemprive Scheduling with Due Dates. *Operations Research*, 27(5), 925-934
41. Sen, T., Sulek, J.M., & Dileepan, P. (2003) Static scheduling research to minimize weighted and unweighted tardiness: A state-of-the-art survey. *International Journal of Production Economics*, 83, 1-12.