

# OBRAZLOŽENJE PRIJEDLOGA TEME DOKTORSKE DISERTACIJE

Kandidat:

Melina Lutvica

Magistar elektrotehnike - diplomirani inženjer elektrotehnike, odsjek računarstvo i informatika

Radni naslov teme doktorske disertacije:

**PSO baziran, podacima vođen algoritam za identifikaciju strukture Bajesove mreže**

## 1 Uvod

Nedostatak jasnog i sveobuhvatnog modela koji kvalitetno opisuje domen predstavlja izazov za upravitelje kompleksnih sistema. Nepotpun prikaz sistema, relevantnih procesa i uzročno posljedičnih relacija između istih, često rezultuje donošenjem neoptimalnih odluka koje mogu biti štetne po sistem. Mnogo je pristupa s kojim se pokušavaju riješiti ili u značajnoj mjeri umanjiti navedeni problemi, a jedan od njih su Bajesove mreže (engl. *Bayesian Network*, *BN*). Bajesove mreže vizualno prikazuju model domena, te prikupljaju informacije iz istog. Ova grafička struktura izabrani domen opisuje i formalizira s ciljem definisanja svih relevantnih entiteta, akcija i kauzalnosti sistema, te omogućava simulaciju vjerovatnih scenarija u istom.

Mnoge su prednosti zbog kojih se, kada je riječ o navedenoj problematici, kao kvalitetan pristup nameću Bajesove mreže. Međutim, bitno je napomenuti da proces kreiranja grafa koji dovoljno dobro modelira realni sistem predstavlja izazov s kojim se susreću i naučnici i domenski stručnjaci. Naime, osamdesetih godina prošlog stoljeća, kada je i uveden termin Bajesovih mreža, razvoj mreže je gotovo isključivo ovisio o domenskom ekspertu i timu ljudi zaduženih za preslikavanje znanja, iskustava i opažanja u oblik grafa. Iako su mogućnosti čovjekovog rezonovanja u mnogim poljima i danas nenadmašene, nesporno je da su u procesu ručne izrade grafa prisutni mnogi nedostaci koji su posljedica ljudske pristrasnosti te ograničene percepcije, zapažanja, memorije i brzine rada.

Dostupnost računarskih kapaciteta te olakšana akvizicija podataka motiviraju sve veći broj komercijalnih organizacija ali i naučnih institucija da vrijeme i resurse investiraju u razvoj automatiziranih metoda i alata za efikasnu i efektivnu analizu prikupljenih podataka. Cilj je da razvijeni alati generišu jasne i upotrebljive informacije na osnovu kojih se mogu donositi odluke. Značajan dio resursa je upravo posvećen razvoju novih te unapređenju postojećih algoritama koji na osnovu podataka koji oslikavaju domen pokušavaju što kvalitetnije konstruisati graf Bajesove mreže. Razvojem ovakvih algoritama, prevazilaze se pomenute slabosti koje u procesu ručne izrade grafa pokazuje čovjek.

Iako se radi o relativno novoj oblasti, naučna zajednica je ponudila razne pristupe rješavanju problema generisanja BN mreže iz podataka. Neki od ponuđenih pristupa su se, kada je riječ o konkretnim problemskim situacijama, pokazali veoma uspješnim. Međutim, generalno govoreći, mnogo je prostora za unapređenje tačnosti, pouzdanosti, performansi te primjenjivosti ovih algoritama, posebno kada je riječ o modeliranju kompleksnih realnih sistema.

U ovom radu će biti analizirani koncepti te objašnjeni algoritmi na kojima se temelje sve značajnije aktuelne metode za generisanje BN mreže iz podataka. Pokušat će se ukazati na mogućnosti unapređenja nekih od prezentiranih pristupa sa naglaskom na potencijal koji u tom smislu nudi primjena meta-heurističkih algoritama. Imajući u vidu da je oblast meta-heurističkih algoritama izuzetno široka, bitno je napomenuti da će u ovom radu fokus biti stavljen na primjenu konkretno meta-heuristike za optimizaciju rojem čestica (engl. *Particle Swarm Optimization, PSO*) koja se pokazala kao vrlo pogodna za primjenu u ovoj oblasti.

Disertacija će pored teorijske analize ponuditi i implementaciju prikazanih koncepata, te analizirati mogućnosti primjene istih na realni problem iz domena upravljanja prirodnim resursima. U saradnji sa ekspertnim timom iz oblasti ekologije i genetike, biti će odabran konkretan problemski domen u smislu geografskog područja, vrste i faktora koji utiču na opstanak vrste. Identifikacijom modela u obliku BN mreže te upotrebom istog za simulaciju i predikciju vjerovatnih događaja, ova disertacija će pokušati doprinijeti kvalitetnijem donošenju odluka za posmatrani domen.

## 2 Pregled stanja u oblasti istraživanja

### 2.1 Uvod

Ovim dokumentom se nastoji dati sveobuhvatan pregled trenutnog stanja u oblasti istraživanja vezano za Bajesove mreže, njihovu definiciju i svojstva, način kreiranja, proces zaključivanja i primjenu u realnim sistemima. Posebna pažnja je posvećena metričkim, podacima vođenim metodama za generisanje grafa, te upotrebi heurističkog algoritma kao mehanizma pretrage u istim. S obzirom da PSO odlikuju odlične performanse kao i relativno jednostavna adaptacija u smislu primjene na datu problematiku, istom će u ovom radu biti posvećena posebna pažnja.

### 2.2 Osnovne karakteristike Bajesovih mreža

#### 2.2.1 Struktura

Bajesova mreža je usmjereni aciklični graf (engl. *Directed Acyclic graph, DAG*) koji strukturirano modelira određeni domenski prostor prikazivanjem domenskih varijabli čvorovima grafa a relacija između datih varijabli usmjerenim granama grafa. Cilj ovakvog modela je da slikovito opiše domen i prikaže raspodjelu združene vjerovatnoće (engl. *joint probability distribution*) za dati domen. Bajesove mreže se zasnivaju na istoimenom teoremu (engl. *Bayes Theorem*)[1] koji se vrlo jednostavno može objasniti s relacijom koja je data u nastavku.

$$P(h|e) = \frac{P(e|h) * P(h)}{P(e)}$$

Naime, prema Bejesu, vjerovatnoća da je hipoteza  $h$  tačna uzimajući u obzir da se pojavio dokaz  $e$  (aposteriori vjerovatnoća), jednaka je normalizovanom umnošku vjerovatnoće pojave dokaza  $e$  u slučaju da je hipoteza  $h$  tačna  $P(e|h)$  (engl. *likelihood*), i vjerovatnoće tačnosti hipoteze  $h$  bez obzira na pojavu dokaza  $e$   $P(h)$  (apriori vjerovatnoća). Normalizacija se vrši dijeljenjem iskaza s vjerovatnoćom pojave dokaza  $e$   $P(e)$ .

Bajesova mreža se zasniva na grafičkoj strukturi koja vizualizira domen i parametrima koji kvantificiraju domen. Graf se koristi da prikaže varijable i veze između njih, a parametri mreže su dati kao skup vrijednosti koje neka varijabla može primiti i vjerovatnoće da će se to u određenim uslovima desiti. Inicijalno stanje mreže se zasniva se na strukturi grafa i apriori vjerovatnoći pojave opisanih stanja. Pojavom novog dokaza se stanje u mreži mijenja provođenjem postupka rezonovanja, interferencije ili učenja. Ovim postupkom se na osnovu Bajesovog teorema računa aposteriori vjerovatnoće, te se ažuriraju vrijednosti zavisnih varijabli u grafu. Drugačije rečeno, Bajesova mreža propagira informaciju u novonastalom dokazu kroz čitav graf i na osnovu toga osvježava stanje u modelu domena [2]. Navedene karakteristike čine Bajesovu mrežu idealnim kandidatom za predstavljanje probalističkog odnosa između definisanih varijabli domena i koriste se za rezonovanje pri postavljanju neizvjesnosti. Iako postoje i druge strukture koje nastoje prikazati raspodjelu združene vjerovatnoće, značaj Bajesovih mreža je upravo u mogućnosti prikazivanja kauzalne prirode domena prikazivanjem i izostavljanjem grane koja povezuje dva čvora, te davanjem usmjerenja datoj grani [3]. Bajesove mreže dati domen modeliraju hijerarhijskom strukturom korištenjem usmjerenih grana koje povezuju čvorove i time prikazuju odnos roditelj – dijete ili prethodnik – nasljednik. Upravo struktura poznata kao Markovljevi pokrivač (engl. *Markov blanket*), koja se za promatrani čvor sastoji od njegovih čvorova roditelja, čvorova djece i njihovih roditeljskih čvorova, podcrtava da bliska okolina promatranog čvora i kauzalna relacija između tih čvorova direktno utiče na stanje posmatranog čvora [4]. Jedna od ključnih karakteristika Bajesovih mreža je i mogućnost prikazivanja indirektnih ovisnosti i uticaja modeliranjem raznih vrsta struktura opisanih pravilom d-separacije (engl. *d-separation*) [5]. Međutim, da bi se graf Bajesove mreže smatrao ispravnim ili potpunim, u datim strukturama ne smije biti petlji, te moraju biti zadovoljeni zahtjevi opisani Markovljevim svojstvom (engl. *Markov property*). Ovo svojstvo nalaže da ne postoje direktne ovisnosti između varijabli modeliranog sistema koje nisu već prikazane granama u grafu Bajesove mreže. Bajesove mreže koje imaju Markovljevo svojstvo se također nazivaju nezavisne mape (engl. *Independence maps*, *I-Maps*) zbog toga što nedostatak grane između čvorova u grafu sugerira neovisnost između datih varijabli sistema. Analogno, zavisnim mapama se nazivaju svi grafovi koji svaku ovisnost između varijabli domena prikazuju postojanjem grane između čvorova koji reprezentiraju varijable. Prilikom modeliranja domena uvijek treba težiti ka kreiranju minimalnog grafa u kojem postojanje grane signalizira postojanje direktne ovisnosti varijabli domena, a nepostojanje grane između varijabli domena signalizira nepostojanje direktne ovisnosti. Ovakvi grafovi, koji su ujedno i zavisne i nezavisne mape, nazivaju se savršeni grafovi jer bi brisanje ili dodavanje bilo koje grane prekršilo data pravila [2][3][6].

### 2.2.2 Parametrizacija

Parametrizacija mreže se vrši definisanjem tzv. tabela uslovne vjerovatnoće (engl. *Conditional Probability Table*, *CPT*) koje kvantificiraju odnos između direktno zavisnih varijabli. Naime, za svaki čvor se definiše skup vrijednosti koje reprezentirana varijabla može primiti, kao tabela uslovne vjerovatnoće za dati čvor. Ova tabela definiše sve moguće kombinacije vrijednosti roditeljskih čvorova, gdje se svaka jedinstvena kombinacija vrijednosti koje uzimaju roditelji naziva instancom roditeljskog skupa. Za svaku instancu roditelja je potrebno definisati vrijednost čvora djeteta i apriori vjerovatnoću da će u datim uslovima posmatrani čvor primiti određenu vrijednost. Uočavamo da je veličina CPT tabele čvora eksponencijalno vezana za broj roditelja, zbog čega je kompaktnost (minimalnost) grafa BN-a izuzetno bitna za tačnost, preglednosti i performanse mreže. Čak u slučaju tzv. Bool-ove mreže postoji  $2^{n+1}$  kombinacija po čvoru, gdje je  $n$  broj roditeljskih čvorova. Ova mreža sa aspekta parametara predstavljaju najjednostavniju varijantu jer modelira domen tako da svaka varijabla može primiti samo jednu od dvije moguće vrijednosti, tačno ili netačno. Imajući u vidu navedene razmjere, jasno je da Bajesove mreže koje se koriste za realne

sisteme mogu vrlo brzo postati izuzetno kompleksne što može dovesti do tzv. prokletstva dimenzija. Ovaj problem još više dolazi do izražaja ako se kreiranje parametarskih tablica vrši manualno.

### 2.2.3 Interferencija

Bajesove mreže na osnovu Bajesovog teorema vrše rezonovanje duž strukture mreže. Ovim postupkom se vrši računanje aposteriori vjerovatnoće i ažuriranje tabela uslovne vjerovatnoće svih čvorova. Ovisno o strukturi mreže i prethodno postavljenim ciljevima, postoji nekoliko vrsta rezonovanja navedenih u literaturi. Dijagnostičko rezonovanje se vrši u svrhe definisanja uzroka na osnovu evidencije simptoma. Ono se vrši u smjeru suprotnom usmjerenju grana i vrlo je često korištena tehnika u oblasti medicine i u dijagnostičkim istraživanjima [7][8][9][10][11][12][13]. Predikciono rezonovanje se vrši u smjeru usmjerenja grana grafa i služi za predviđanje vjerovatnih posljedica na osnovu zabilježenih dokaza. Ovakvim rezonovanjem se može vršiti i simuliranje scenarija značajnih za poznavaoce domena. Postoji veliki broj publikacija primijene ove tehnike u oblasti epidemiologije i ekologije [14][15][16][17][18]. Interkauzalno rezonovanje se koristi za zaključivanje o uzroku nekog događaja. Ovaj proces metodom eliminacije može zaključiti o vjerovatnom uzroku događaja između više mogućih uzroka (engl. *Explaining away*). Kombinirano rezonovanje predstavlja oblik rezonovanja gdje se s obzirom na pojavu dokaza i postavljene ciljeve, ne može jasno definisati smjer niti proces rezonovanja. Pored navedenog, Bajesove mreže se vrlo često koriste i kao klasifikatori gdje, čak i u svojoj najjednostavnijoj formi (tzv. naivne Bajesove mreže), postižu jako dobre performanse [19][20][21].

Neovisno o kojem tipu rezonovanja se radi, postoje dvije osnovne grupe algoritama koji se koriste za zaključivanje – egzaktni i aproksimirani [2]. U posljednje tri decenije je definisan veliki broj ovih algoritama, od kojih su najpoznatiji dostupni u skoro svim softverskim paketima za rad s Bajesovim mrežama. Kompleksnost primjene algoritma raste sa kompleksnošću strukture mreže. Na mreže koje su jednostavni lanci čvorova je dovoljno iterativno primjenjivati Bajesov teorem. Zaključivanje u jednostavnim strukturama stabla se vrši pomoću lokalnih proračuna i algoritma za prosljeđivanje poruka između čvorova. U oba slučaja je moguće koristiti egzaktnu algoritmu. Međutim, kada su parovi čvorova povezani s više staza, algoritam za zaključivanje postaju složeniji. Za neke mreže zaključivanje s egzaktnim algoritmima tad postaje računski neizvodivo, pa se u tim slučajevima pribjegava aproksimiranim algoritmima za zaključivanje. Teorijski gledano, obje vrste algoritama spadaju u računski kompleksne algoritme NP-teške kategorije. Međutim, u praksi, brzina zaključivanja zavisi od faktora kao što su struktura mreže, povezanost, broj neusmjerenih petlji i drugih strukturnih komponenti koji utiču na performanse algoritma. Postoje različiti algoritmi koji su se u pokazali vrlo efikasnim za određeni podskup struktura, problemskih prostora i tipova mreže.

Egzaktni algoritmi se mogu svrstati u neku od podklasa od kojih su najpoznatije „Poly-tree“, klastering, algoritmi za uvjetovanje, zamjenu grana, eliminaciju, simbolički algoritmi te algoritmi s metodom diferencijala [22]. Svaka od ovih klasa poznaje nekoliko značajnih algoritama sa raznim varijacijama, hibridima, generalizacijama i heurističkim implementacijama. Od svih njih najpoznatiji je *Junction tree* ili još poznat kao *Clique tree* algoritam. Ovaj algoritam se obično primjenjuje na probleme NP-potpune klase kompleksnosti zbog čega se u praksi efikasno primjenjuje na mreže sa do tridesetak čvorova [2].

Aproksimirani algoritmi se dijele na algoritme stohastičke simulacije, metode simplifikacije modela, te metode bazirane na pretraživanju i propagaciji vjerovatnoće petljama. Iako postoje mnoge različite implementacije ovih algoritama, najpoznatije spadaju u grupu Markovljev lanac Monte Carlo (engl. *Markov Chain Monte Carlo*, *MCMCM*) algoritama [22]. U posljednjih par godina pojavilo se i nekoliko

novih algoritama koji su uglavnom hibridi postojećih ideja namijenjeni za efikasno rješavanje određene klase problema procesa zaključivanja [23][24][25][26][27].

#### 2.2.4 Klasifikacija i aplikacija

Iako se u literaturi najčešće spominje generalni oblik Bajesovih mreža, zbog različitih potreba, vremenom su razvijene specifične klase Bajesovih mreža koje implementiraju određene posebnosti. Primjer takvih Bajesovih mreža predstavljaju Bajesove mreže za donošenje odluka (engl.*Bayesian Decision Networks, BDN*). Prilikom odlučivanja je važna mogućnost uključivanja preferencija u proces izbora neke od dostupnih akcija. U ove svrhe se koristi teorija korisnosti (engl.*Utility theory*) koja omogućava da se vrši zaključivanje s preferencijama. Ova klasa BN mreže definiše dodatne vrste čvorova koji oslikavaju situaciju donošenja odluke optimiziranjem ranije definisane funkcije cilja koja ovisi o vjerovatnoći da neka domenska varijabla primi neku od mogućih vrijednosti. Definisana funkcija cilja kvantificira preferencije, odražavajući "korisnost" ishoda, tako što mapira preferencije na realne brojeve. Ovakve mreže nalaze veliki broj primjena posebno u domenu upravljanja i planiranja prirodnim resursima [2][28][29].

Dodavanjem vremenske dimenzije Bajesovoj mreži nastaje tzv. dinamična Bajesova mreža (engl.*Dynamic Bayesian Network, DBN*). Ovakve mreže omogućavaju eksplicitno modeliranje i rezonovanje kroz vrijeme [2][30][31]. Kombiniranjem svojstava mreža za donošenje odluka i dinamičkih mreža nastaju dinamičke Bajesove mreže za donošenje odluka (engl.*Dynamic Decision Network, DDN*). Ovakve mreže su pogodne za prikazivanje složenih problema planiranja i sekvencijalnog donošenja odluka.

Jedna od slabije istraženih oblasti su objektno-orijentirane Bajesove mreže (engl.*Object Oriented Bayesian Network, OOBN*). Ova generalizacija Bajesove mreže je vrlo pogodna za izrazito kompleksne mreže koje predstavljaju realne sisteme. One se temelje na podjeli problemskog prostora na više domena prikazanih hijerarhijskom strukturom komponenti koje su opet predstavljene Bajesovim mrežama [2][32][33]. Ovakvi primjeri se u literaturi ipak susreću rjeđe, posebno iz razloga što Bajesove mreže još uvijek nemaju širokopojasnu primjenu u komercijalnim sistemima.

Bajesovih mreže su primijenjene u mnogim oblastima, pri čemu je najveći broj publikacija i članaka posvećen problemima iz oblasti medicine, tačnije dijagnostike i epidemiologije. Drugo mjesto po broju publikacija zauzima oblast za upravljanje prirodnim resursima i ekologiju u najširem smislu. Na trećem mjestu su igre i poslovne aplikacije koje u posljednjih deset godina pokazuju sve više zanimanja za upotrebu Bajesovih mreža u svrhe predikcije i odlučivanja. S obzirom da je interes ove disertacije na primjeni Bajesovih mreža u realnim sistemima za upravljanje prirodnim resursima, s posebnim su zanimanjem proučene publikacije [34][35][36][37][38][39][40][41][42] koje potvrđuju da Bajesove mreže daju odlične rezultate za donošenje odluka u sredinama s visokim stepenom neizvjesnosti.

Gotovo sve navedene klase Bajesovih mreža su dostupne u softverskim paketima koji su ciljano razvijeni za rad s ovim strukturama. U literaturi se najčešće spominju AgenaRisk[43], BayesiaLab[44], BNT [45], BUGS [2], CABeN[2], CaMML[46], Hugin [2], IDEAL [47], Netica[48], GeNIe/SMILE [49] i TETRAD [50].

### 2.3 Podacima vođeni algoritmi za generisanje grafa Bajesove mreže

Iako sam proces kreiranja Bajesove mreže nije u potpunosti formaliziran, u literaturi se često spominje proces pod nazivom inženjering znanja za Bajesove mreže (engl.*Knowledge Engineering with Bayesian Networks, KEBN*) [2][51]. KEBN podržava metodologije vodopada i spiralnog pristupa, u smislu da za obje varijante predviđa faze razvoja mreže, izbora varijabli, definisanja strukture grafa, parametrizaciju tabela, treniranje mreže itd. Tokom faze definisanja strukture grafa i faze definisanja parametara mreže se može

koristiti manualni pristup kao i automatizirani alati. U posljednjoj deceniji se posebna pažnja posvećuje integraciji ova dva pristupa, tako da se nastoji znanja domenskog eksperta ugraditi u automatizirane alate za izgradnju grafa. Potreba za automatizacijom procesa je posebno naglašena kada je domen kompleksan, te kada su graf i broj mogućih struktura izrazito veliki. Stoga je vremenom razvijena oblast koja se bavi izučavanjem algoritama za učenje strukture i parametara Bajesove mreže iz podataka. Proces prepoznavanja Bajesove mreže iz podataka poznaje četiri različita slučaja [52]:

- poznata struktura i potpuni podaci,
- nepoznata struktura i potpuni podaci,
- poznata struktura i nepotpuni podaci i
- nepoznata struktura i nepotpuni podaci

U posljednjih nekoliko decenija je razvijen veliki broj algoritama koji su specijalizirani za neki od navedenih slučajeva. U ovoj disertaciji će fokus biti na algoritmima za prepoznavanja strukture iz potpunih podataka, te će samo ukratko biti govora ostalim slučajevima i algoritmima iz datih oblasti.

Slučaj poznate strukture i potpunih podataka se koristi za parametrizaciju modela BN-a. Ovo predstavlja jedno od najbolje istraženih područja kada je riječ o podacima vođenom učenju Bajesove mreže. S obzirom da je kreiranje tabela uslovne vjerovatnoće zadatak koji je eksponencijalne kompleksnosti u odnosu na broj čvorova, korištenje automatiziranog procesa u ove svrhe može biti od velikog značaja. Multinominalna parametrizacija ili tzv. Spiegelhalter i Lauritzen metoda [53] je postala opće-prihvaćen algoritam koji je implementiran i u mnogim softverskim paketima. MLE (engl.*Maximum likelihood estimation*) je također vrlo rasprostranjen algoritam [2] korišten u ove svrhe. U nedavno publiciranoj literaturi se mogu naći i neke druge generalizacije ili varijante algoritma za parametrizaciju mreže [54][55][56][57][58].

Problemu poznate strukture i nepotpunih podataka se posvetilo nekoliko zanimljivih publikacija, kao one posvećene algoritmu maksimizacije očekivanja - EM (engl.*Expectation maximization*) [2][59] i algoritmu pod nazivom *Gibbs Sampling* [2][59] koji je jedan od najznačajnijih algoritma iz skupine Markovljev lanac Monte Carlo metoda. Oba algoritma se baziraju na strategiji dopunjavanja nedostajućih podataka na osnovu postojećih podataka. Problem ovog pristupa je velika mogućnost greške i pristranosti podataka kada je samo manji broj podataka dostupan i poznat. Kao poboljšan koncept se pojavio BC algoritam (engl.*Bound and Collapse*) [60] koji predlaže determinističku metodu za aproksimaciju varijance aposteriori distribucije za definisanje uslovnih vjerovatnoća iz nepotpunih podataka.

Najteži problem, slučaj nepoznate strukture i nepotpunih podataka je u odnosu na ostale vrlo malo istražen. Jedan od algoritama predloženih za ovaj slučaj je modifikacija EM algoritma nazvana SEM, ili strukturalni EM (engl. *Structural Expectation maximization*) [61]. Algoritam kombinira standardni EM algoritam, koji optimizira parametre mreže s pretraživanjem struktura za izbor modela. Drugi algoritam koji je predložen u ove svrhe je varijanta BC algoritma sa dodatkom za prepoznavanje strukture.

### 2.3.1 Nepoznata struktura i potpuni podaci

Slučaj nepoznate strukture i potpunih podataka je u fokusu ove disertacije i stoga detaljnije opisan u ovom radu. Ovi algoritmi su tema velikog broja aktualnih istraživanja. Još od kraja 90-tih pa do danas postoje značajne publikacije koje proučavaju različite mehanizme iz oblasti statistike, teorije informacija, heuristike i mnogih drugih oblasti u cilju definisanja efikasnog i efektivnog algoritma za prepoznavanje strukture Bajesove mreže. Iako publikacije [62][63][64] uvode različite algoritme i sistematizacije za iste, ipak se svi mogu svrstati u jedan od dva glavna pristupa rješavanja: metode bazirane na testovima uslovne

neovisnosti (engl.*Conditional Independence Tests*) i metode bazirane na ciljnoj funkciji i pretraživanju (metrički algoritmi). U literaturi se spominju i hibridni algoritmi koji kombinuju ova dva pristupa.

#### 2.3.1.1 Algoritmi bazirani na testovima nezavisnosti

Ove metode vrše kvalitativnu analizu zavisnosti i nezavisnosti između varijabli u podacima i pokušavaju da što više tih relacija predstave u obliku grafa. Ovim testom se izračunava da li postoji statistički značajna povezanost između hipoteze predstavljene potencijalnim grafom, i opažene frekvencije predstavljene bazom podataka. Neki od značajnijih algoritama su PC (engl.*Power Constructor*) [65], BNPC (engl.*Bayesian Network Power Constructor*) [66] i algoritmi bazirani na d-separaciji [67].

#### 2.3.1.2 Metode bazirane na ciljnoj funkciji (metrički algoritmi)

Ovakvi algoritmi pristupaju problemu definisanja strukture mreže kao optimizacijskom problemu. Takav problem se može opisati kao problem izbora optimalne alternative iz skupa dostupnih alternativa kako bi se postigao neki cilj. Realni problemi optimizacije su najčešće vrlo kompleksni problemi koji su multi-dimenzionalni i višekriterijalni te gotovo uvijek imaju ograničenja. Metode bazirane na ciljnoj funkciji i pretraživanju (tzv. metričke metode) upravo vrše pretragu prostora za onom reprezentacijom domena koja maksimizira funkciju cilja koja opisuje podudarnost grafičke strukture i podataka. S obzirom da ovakvi problemi zbog kompleksnosti prostora koji treba pretražiti spadaju u NP-teške optimizacijske probleme, upotreba heurističkih algoritama u ove svrhe pokazuje odlične rezultate.

Heuristički algoritam je algoritam koji pronalazi rješenje kompleksnog problema ili pojednostavljenog modela kompleksnog problema, stavljajući pri tome dokazivost egzistencije, tačnosti i optimalnosti rješenja u drugi plan. Tokom pretraživanja problemskog prostora, heuristički algoritam se vodi neformalnim mehanizmima kao što su iskustvo, znanje ili intuicija koji su vrlo često preslikani iz prirodnih procesa. Svaki heuristički algoritam ili konkretna implementacija koristi svoj vlastiti skup mehanizama i procedura s kojim nastoji da što efikasnije dođe do optimalnog rješenja. Generalno govoreći, većina heurističkih algoritama se konceptualno može uprošteno predstaviti u obliku dijagrama toka na sljedeći način. Prvo se vrši izbor odnosno generisanje jednog ili više početnih rješenja. Nadalje, sve dok uslov zaustavljanja nije zadovoljen, vrši se evaluacija kvaliteta trenutnih rješenja, izbora novih, kvalitetnijih potencijalnih rješenja, te eventualna zamjena tekućeg rješenja s kvalitetnijim rješenjima. Algoritam se zaustavlja i najbolje odnosno najbolja rješenja se proglašavaju konačnim tek po ispunjenju jednog ili više uslova zaustavljanja. Iz opisa opće procedure heurističkog algoritma je vidljivo da svaki heuristički algoritam posjeduje sljedeće bitne elemente [68][69]:

- jedna ili više ciljnih funkcija koje se optimiziraju,
- mehanizam izbora početnog rješenja,
- memorija potencijalnih rješenja,
- mehanizam procjene kvalitete tekućeg rješenja,
- mehanizam transformacije trenutnog rješenja / mehanizam izbora novog rješenja,
- uslovi zaustavljanja algoritma.

Metrički algoritmi kombiniraju i implementiraju date elemente u svrhu definisanja strukture grafa mreže koja najbolje modelira domenske podatke. Konkretni metrički algoritam je definisan algoritmom pretraživanja koji koristi, funkcijom cilja koju optimizira i načinom na koji predstavlja problemski prostor. U većini slučajeva algoritam pretraživanja je baziran na lokalnom pretraživanju, tabu pretraživanju ili drugoj heurističkoj metodi pretraživanja. Problemski prostor može biti ograničen i neograničen, te se mogu koristiti različite metode za predstavljanje potencijalnog rješenja, tj. grafa, pri čemu je najčešće

korištena matrica povezanosti čvorova. Definisana funkcija cilja najčešće predstavlja neku funkciju koja opisuje mjeru fitovanja između grafa i podataka. Ista se može bazirati na Bajesovim principima kao Bajesova i Dirihletova funkcija raspodjele, ili na principima teorije informacija kao funkcija minimalne dužine poruke (engl. *Minimum Description Length*) ili logaritamske vjerovatnoće (engl. *Log Likelihood*, *LL*). U nastavku će biti pobrojani samo neki od poznatijih algoritama.

BDeu (engl. *Bayesian-Dirichlet likelihood equivalence uniform joint distribution*) [70] je metrički algoritam koji koristi lokalno ili tabu pretraživanje uz funkciju cilja temeljenu na Bajesovom teoremu.

K2 algoritam [71] je jedan od prvih algoritama za rekonstrukciju mreže bez ograničenja i koristi sličnu funkciju cilja temeljenu na Bajesovoj teoriji. S obzirom da ovaj algoritam u originalnom obliku ima eksponencijalnu kompleksnost u odnosu na broj varijabli, predloženo je uređivanje redoslijeda varijabli kako bi se postigle bolje performanse algoritma. Samo uređivanje redoslijeda može vršiti domenski ekspert ili drugi algoritam, što uvodi dodatno računanje, međutim rješava problem usmjerenja grana.

Kutato algoritam [72] je metrički algoritam koji koristi entropiju kao mjeru podudarnosti grafa i podataka. Vrlo često su korištene i funkcije bazirane na principu najkraće poruke (MDL) [73][74] koje su inspirisane Šenanovom mjerom informacija (engl. *Claude Shannon - Measure of information*). Cilj ovog pristupa je naći optimalnu mjeru između jednostavnosti modela i podudarnosti s podacima, tako što minimizira dužinu opisa modela i podataka, krećući od pretpostavke da je model ispravan. MML (engl. *Minimum Message Length*) predstavlja ideološki vrlo sličnu tehniku, orijentisanu ka Bajesovim principima.

S obzirom da neki od navedenih algoritama ne vode računa o usmjerenju grana i kauzalnoj prirodi mreže, razvijeno je nekoliko varijacija ovih algoritama. CaMML (engl. *Casual Discovery via MML*) [2][75] i Hibridni CaMML (engl. *Hybrid Causal Discovery via MML*) [2], su samo neki od algoritama koji uvode dodatne mehanizme za uključivanje usmjerenja grana u formiranje mreže. Detaljan pregled raznih pristupa za prepoznavanje kauzalnih mreža dat je i u [76][77].

Jedan od najpoznatijih hibridnih algoritama je BENEDICT [78] koji se temelji na mjerenju razlike između rezultata testova uslovne nezavisnosti na kandidatskoj mreži i rezultata dobivenih na bazi podataka. Što je ova razlika manja to mreža bolje odgovara podacima. Agregacija svih lokalnih neslaganja rezultira mjerom globalnog raskoraka između mreže i baze podataka, što zapravo predstavlja funkciju cilja koja se nastoji minimizirati.

U literaturi se spominju i algoritmi koji koriste sofisticirane meta-heurističke algoritme. Najčešće su korišteni algoritmi kao „*Sparse Candidate Algorithm*“ [79], algoritam na bazi uređivanja rasporeda [80][81], algoritam najkraćeg puta [82], heuristički „*Equivalent search*“ algoritam [83]. U najnovijim publikacijama prisutni su algoritmi paralelnog pretraživanja [84], optimizacije kolonijom mrava [85], genetički algoritmi [86][87], BOA (engl. *Bayesian Optimization Algorithm*) [88] i evolucione strategije [25]. U većini ovih publikacija je riječ o modifikaciji nekog od poznatih meta-heurističkih algoritama u cilju prilagođavanja datom problemskom domenu.

Za ovu disertaciju je posebno zanimljiva heuristika optimizacije rojem čestica, te će u zasebnom poglavlju biti dat pregled literature i baznih koncepata datog meta-heurističkog algoritma koji je vrlo pogodan za upotrebu u metričkim metodama generisanja grafa Bajesove mreže.

### 2.3.2 Mjere performansi algoritama za generisanje grafa

S obzirom da ovo istraživanje treba da uporedi različite podacima vođene algoritme za generisanje grafa BN-a, mjerenje performansi istih je značajan alat za utvrđivanje kvaliteta algoritama. U sklopu istraživanja je izvršen pregled literature [2][62] i zabilježene su najznačajnije mjere performansi koje će biti korištene u daljem istraživanju. Neke se mogu koristiti za utvrđivanje performansi algoritma kad je traženo rješenje

već poznato(kao npr. na *benchmark* problemima), a druge kada se radi o kompleksnim problemima iz realnog sektora gdje nije unaprijed poznato traženo rješenje.

KL distanca (engl. *Kullback–Leibler distance*) je mjera koja definiše distancu između distribucije vjerovatnoće asocirane sa bazom podataka i distribucije vjerovatnoće asocirane s generisanim grafom Bajesove mreže. Ovom mjerom moguće je izračunati koliko dobro identificirani graf aproksimira dostupne podatke. S obzirom da računanje ovakve udaljenosti ima eksponencijalnu kompleksnost, pogodno je koristiti monotono-opadajuću transformaciju date funkcije [89]. Vrijednost dobivena ovom transformacijom je direktno proporcionalna sa mjerom podudarnosti modela i podataka. Bitno je samo obratiti pažnju da vrlo visoke vrijednosti ove mjere mogu također uputiti na prezasićenje (engl. *over-fitting*).

Kauzalna KL distanca (engl. *Causal Kullback–Leibler distance*) je mjera koja ne tretira samo jedan model kao ispravan, nego sve Markovljeve ekvivalentne modele. Na ovaj način CKL jedinstveno uvažava univerzalnost strukture i tačnost parametara, u cilju definisanja najpogodnije mjere validnosti modela kauzalnih mreža. Mnogi eksperimenti su vršeni korištenjem ove mjere kao što je navedeno u [90].

Faktor podudarnosti prezentovan u [91] poredi sličnost prethodno defnisanog „ispravnog“ rješenja i generiranog grafa jednostavnom ali sveobuhvatnom relacijom koja zbog ograničenja koja nameće format ovog rada ovdje neće biti detaljnije analizirana.

Bajesove metrike kao BDeu [68], K2 [69] i BIC [92] mjere kvalitet modela korištenjem distribucije vjerovatnoće. K2 i BDeu se samo razlikuju u izboru parametara u Dirihletovoj distribuciji, dok BIC pored mjere za vjerovatnost mreže koristi i penalnu funkciju kompleksnosti mreže. Sve ove metrike se interpretiraju na jednak način. Što je veća vrijednost metrike, to je aproksimacija podataka grafom bolja.

Vjerovatnoća pojave dokaza je jedna od mjera koja se primjenjuje tokom procesa zaključivanja u mreži. Naime, u trenutku kada se pojavi novi dokaz prema kojem se ažuriraju vrijednosti vjerovatnoća u mreži, situacija da je apriori vjerovatnoća pojave datog dokaza jednaka nuli nije moguća i time se može utvrditi kontradiktornost između dokaza i modela. Naravno, na ekspertu je da procjeni da li je dati dokaz relevantan, no ukoliko jeste, jasno je da model nije validan. Iz tog razloga je definirana ova mjera konflikta između dokaza i modela [93], koja, ukoliko je pozitivna, indicira da postoji konflikt, a sa porastom ove mjere raste i odstupanje između modela i dokaza.

## 2.4 Optimizacija rojem čestica

Optimizacija rojem čestica(engl.*Particle Swarm Optimization, PSO*)je meta-heuristički algoritam za optimizaciju koji je inspirisan prirodnim procesima unutra roja čestica. Ovaj algoritam, baziran na principima inteligencije jata ili roja nastoji oponašati roj pčela u potrazi za hranom kakav se susreće u prirodi. Algoritam kreira skup čestica u problemskom prostoru koje čine inicijalnu populaciju, pri čemu svaka jedinka reprezentira potencijalno rješenje datog problema. Svaka čestica vlastiti pravac kretanja kroz problemski prostor (pomak)određuje na osnovu tri aspekta (vektora): aspekt vlastite trenutne lokacije, aspekt historijski najbolje vlastite lokacije, i aspekta lokacije jedne ili više najboljih čestica roja(informant). Na pomak čestice također utiče faktor slučajnosti koji sprječava preuranjenu konvergenciju algoritma. Potom se evaluira fitnes čestice na novoj lokaciji, te se na osnovu najboljeg fitnesa u roju bira čestica informant, ili više njih. Ovaj postupak se ponavlja za sve čestice u roju dok se ne ispune uslovi zaustavljanja [94]. S obzirom da je kretanje svake čestice uvjetovano i kretanjem drugih,

najboljih, čestica, ali i slučajnim izborom nefavorizirane lokacije, PSO algoritam osigurava stabilnu konvergenciju većine populacije. Uticaj čestica jednih na druge se često karakteriše kao faktor socijalnog umrežavanja koji uspostavlja grupnu inteligenciju. Svi faktori koji učestvuju u određivanju nove lokacije čestice objedinjeni su u formuli koja oslikava osnovnu ideju ove heuristike.

$$v_i^j(k+1) = c_1^1 \cdot v_i^j(k) + c_1^2 \cdot (x_i^{j*}(k) - x_i^j(k)) + c_1^3 \cdot (x_i^{r*}(k) - x_i^j(k))$$

PSO je jedan od najzapaženijih algoritama iz skupa primjerski baziranih (engl. *Instance based*) heuristika koja generiše nova rješenja koristeći se skupom postojećih rješenja. Iz tog razloga se po pitanju performansi često poredi s drugim algoritmima iz ove grupe, kao što su genetički algoritam i evolucione strategije. Jednostavnost koncepta, pogodnost za implementaciju u raznim softverskim paketima i mali broj parametara koji zahtijevaju konfiguraciju su značajne prednosti PSO algoritma i uveliko zaslužne za njegovu široku adaptaciju u različitim oblastima. Navedeni algoritam predstavlja samo osnovnu ideju koja je godinama predmet mnogih istraživanja. Data istraživanja su dovela do dvije vrste unaprjeđenja PSO algoritma, varijacije baznog PSO algoritma i modifikacije PSO algoritma.

#### 2.4.1 Varijacije baznog PSO algoritma

Od 2006. godine PSO zajednica je do sad predstavila tri standardne verzije PSO algoritma, SPSO 2006, 2007, i 2011 [95]. Iako su sve ove varijante bazirane na istim osnovnim principima ove meta-heuristike, formulacija i implementacija baznih elemenata, te njihova aplikacija na problemski domen se razlikuje kao što je prikazano u brojim publikacijama na ovu temu [96][97][98][99][100][101][102][103][104].

Problemski prostor (engl. *Search Space*) predstavlja prostor potencijalnih rješenja. Iako je inicijalno PSO heuristika bila okrenuta kontinualnim problemima, postoje i varijante PSO-a koje se koriste za rješavanje problema u diskretnom prostoru. Neke varijante za rješavanje diskretnih problema su pristup zaokruživanja, binarni PSO, trinarni PSO, pristup kaznene funkcije, pristup prilagodbe funkcije, pristup skupa, hibridni pristup i drugi predstavljeni u [100].

Jato ili stado (engl. *Swarm*) predstavlja populaciju ili skup potencijalnih rješenja. Iako su prvobitne varijante implementirale adaptivne mehanizme za računanje broja čestica u populaciji, kasnije je teorijskim putem utvrđeno da je poželjna veličina populacije najčešće između trideset i četrdeset čestica, te su sve češće varijante sa konstantnim brojem čestica [97].

Inicijalizacija čestica (engl. *Particle Initialization*) predstavlja način izbora početnog rješenja ili početne lokacije za svaku česticu pri inicijalizaciji jata. Čestica predstavlja individuu koja je dio jata, ili drugačije rečeno, potencijalno rješenje. Inicijalno stanje se određuje slučajnim izborom ili nekim definisanim algoritmom koji može, a ne mora, biti heuristički algoritam. Mehanizam inicijalizacije je moguće implementirati tako da se izborom inicijalnog rješenja utiče na ujednačenost pokrivenosti problemskog prostora.

Vektor brzine (engl. *Velocity*) predstavlja vektor koji prikazuje pravac kretanja čestice. Iako standardni PSO ne uvodi ograničenja brzine, poznata je i varijanta PSO algoritma [103] koja nastoji ograničiti brzinu kako bi se uvela kontrola nad kretanju čestice i istraživanju prostora. Na ovaj način se pretraga može ograničiti na lokalnu ili proširiti na globalnu pretragu. Ovaj mehanizam se uglavnom implementira kao funkcija vremena ili broja iteracija koja pretraživanje od globalnog postepeno vodi ka lokalnom.

Faktor inercije  $W$  (engl. *Inertia weight*) je komponenta vektora brzine koja kontroliše uticaj historijski najboljih rezultata na trenutna rješenja. Ova komponenta nije definisana izvornom varijantom PSO algoritma, već je ista u [105] predložena i opravdana jer dokazano poboljšava performanse algoritma. Vremenom je uvrštena u standardni algoritam PSO-a. Ona je jedna od rijetkih veličina PSO-a koju je potrebno konfigurisati. S obzirom da je nekoliko istraživanja posvećeno definisanju optimalne vrijednosti ove veličine, vremenom je utvrđen preporučeni opseg ove i drugih konfigurabilnih veličina PSO-a. S obzirom da su preporučene vrijednosti pokazale dobre rezultate, ova veličina se često postavlja kao konstantna vrijednost ili interval vrijednosti [98]. Također postoje i varijante PSO-a u kojima se nastoji dinamički odrediti ova veličina [99].

Faktori učenja (engl. *Learning factor*) predstavljaju naklonost čestice prema vlastitom uspjehu ili prema uspjehu okoline, označene kao  $C1$  ili kognitivni faktor učenja i  $C2$  ili socijalni faktor učenja, respektivno. Obje vrijednosti se uglavnom definišu kao konstante u intervalu između minimalne i maksimalne predviđene vrijednosti. Inicijalna varijanta PSO algoritma predlaže uključivanje faktora učenja u kalkulaciju vektora brzine. Međutim, istraživači u [102] su za određeni problemski prostor pokazali da ove konstante ne utiču na poboljšanje performanse algoritma, te je i varijanta PSO-a bez ovih faktora zastupljena u literaturi.

Vođa ili informant (engl. *Leader, Informant*) predstavlja česticu koja zbog dobrog fitnesa upućuje druge čestice u prema lokaciji problemskog prostora koji daje dobre rezultate. Ovisno o tome kako se bira informant, prepoznato je nekoliko varijanti PSO algoritma. Ukoliko između čestica ne postoji razmjena informacija, algoritam poprima karakteristike nevođene pretrage. U ovom slučaju je socijalna komponenta jednaka nuli i pretraživanje je vođeno samo vlastitom najboljom performansom,  $Pbest$  (engl. *Personal Best*). U slučaju da se informant bira kao najbolja čestica iz skupa okolnih čestica onda se govori o lokalno najboljoj čestici ili  $Lbest$  (engl. *Local Best*). Sama okolina se može unaprijed definisati ili može biti adaptivna. Globalno najbolji ili  $Gbest$  (engl. *Global Best*) je varijanta PSO-a u kojoj se informanti biraju iz skupa svih čestica jata [106].

Topologija okoline (engl. *Neighborhood topology*) definiše skup čestica koje formiraju okolinu date čestice i način izbora istih. Topologija je direktno vezana i za način izbora jednog ili više informanta i može značajno uticati na performanse algoritma, posebno u pogledu pronalaska globalnog optimuma [107]. U literaturi se prepoznaju varijante sa tzv. nepovezanim topologijom gdje čestice nisu povezane, te ne postoji razmjena informacija [108]. Topologija prstena je jedna od najjednostavnijih topologija gdje čestica razmjenjuje informacije samo s dva direktna susjeda [108]. Topologija zvijezde je topologija gdje sve čestice dobivaju informacije samo od jednog informanta koji se bira na nivou čitave populacije [108]. Topologija drveta prepoznaje više informanata i informacija se prenosi kroz hijerarhijski definiranu strukturu [109]. Potpuno povezana topologija povezuje svaku česticu sa svim ostalim česticama jata [108]. Među novijim topologijama se nalazi i topologija dinamičke okoline u kojoj se informant bira iz okoline čestice koja se dinamički određuje tokom izvršenja algoritma [110]. Kompleksnije topologije osiguravaju bolju povezanost čestica i prijenos informacija kroz populaciju, ali isto tako mogu negativno da utiču na performanse algoritma i računarske resurse koji su potrebni za pronalazak rješenja.

SPSO 2011 je posljednja varijanta standardnog PSO algoritma za koju se smatra da objedinjuje mnoge elemente koje performanse ovog algoritma značajno unaprjeđuju. Ova varijanta se pokazala izuzetno efikasnom za optimizaciju unimodalnih i separabilnih funkcija, sa brзом konvergencijom ka globalnom optimumu. Za neke rotirane uninodalne funkcije su ustanovljene dobre performanse, dok su performanse za neke multimodalne i izuzetno kompleksne kompozicijske funkcije bile slabije [96]. Ova

varijanta algoritma se također može koristiti kao *benchmark* za poređenje performansi s drugim PSO varijantama ili heurističkim algoritmima.

#### 2.4.2 Modifikacije PSO algoritma

Mogućnost primjene PSO-a na širok dijapazon problemskih situacija je potakla razvoj nekoliko modifikacija izvornog PSO algoritma. U nastavku su navedene najznačajnije modifikacije PSO-a. U sklopu ovih modifikacija postoji veliki broj varijanti PSO-a koje dijele manje razlike u implementaciji, pri čemu svaka od njih nastoji da poboljša performanse algoritma doradom nekog od mehanizama.

*PSO za pronalazak jedinstvenog rješenja* je izvorni i ujedno najjednostavniji oblik PSO-a. Koristi se za pronalazak jedinstvenog rješenja jednokriterijalnih, kontinualnih, neograničenih, statičkih optimizacijskih problema.

*PSO za pronalazak više rješenja.* Ova modifikacija PSO-a uvedena je za rješavanje višekriterijalnih optimizacijskih problema, te se u literaturi zato često naziva MOPSO (engl. *Multi-Objective Particle Swarm Optimization*). Cilj ove modifikacije je da maksimizira broj pronađenih elemenata pareto fronta [111], da minimizira distancu između pronađenog pareto fronta i globalnog pareto fronta za neki problem, te da maksimizira pokrivenost pareto fronta tako da su pronađena rješenja uniformno distribuirana. U svrhu postizanja navedenih ciljeva razvijeno je nekoliko pristupa rješavanju problema višekriterijalne optimizacije PSO-om [112]. Pristup agregacije kombinira sve kriterije u jedan i time višekriterijalni problem aproksimira jednokriterijalnim problemom. U tom slučaju se agregirana ciljna funkcija optimizira nekom od postojećih varijanti PSO-a. Pristup više populacija koristi više sub-populacija od kojih svaka nastoji optimizirati jedan od kriterija, te se pronađena rješenja zatim prikupljaju i ogovarajuća izabiru kao konačna rješenja. Pareto bazirani pristup modificira topologiju i izbor informanata kako bi se osigurao izbor nedominiranog informanta (potencijalnog rješenja) u odnosu na čitavo jato. Kombinirani pristup koristi kombinaciju nekih od navedenih pristupa u svrhe poboljšanja performansi MOPSO-a. U literaturi se mogu pronaći i drugi pristupi koji se ne mogu jasno svrstati u neku od navedenih kategorija [113][114][115][116][117].

*PSO za optimizaciju s ograničenjima.* Standardni PSO vrši pretraživanje na cijelom problemskom prostoru, dok ova modifikacija implementira posebne mjere kako bi se uzela u obzir ograničenja problemskog prostora. Na ovaj način se smanjuje prostor i broj dozvoljenih rješenja. U literaturi se može naći nekoliko različitih pristupa uključivanja ograničenja u optimizacijski problem. Stoga ova modifikacija PSO-a poznaje različite varijante, od onemogućavanja izbora nedozvoljenog rješenja ili odbacivanja istog, do upotrebe penalne funkcije kako bi se nedozvoljeno potencijalno rješenje učinilo neoptimalnim [118][119].

*PSO za dinamička okruženja.* Ova modifikacija PSO algoritma je razvijena za probleme koji su promjenjivi u vremenu. PSO nastoji korištenjem standardnih procedura pronaći optimalno rješenje prije no što dođe do iduće promjene u okruženju [103].

*Diskretni PSO* je modifikacija PSO algoritma koja prilagođava PSO za pretraživanje u diskretiziranom prostoru. Ovo se najčešće postiže uvođenjem posebne reprezentacije potencijalnog rješenja i prilagodbom mehanizma za računanje pomaka datom zapisu [103]. Kao što je ranije spomenuto, ova modifikacija poznaje različite varijante od kojih je binarni PSO jedan od poznatijih.

*Paralelni PSO* je modifikacija PSO-a koja implementira mehanizme za paralelno pretraživanje prostora korištenjem više populacija. Ovisno o tome da li se rješava jednokriterijalni ili višekriterijalni problem,

postoje različite varijante ove modifikacije. U nekim su jata povezana i služe za optimizaciju istog kriterija ali pretražuju odvojene oblasti problemskog prostora. U drugima su jata u potpunosti odvojena i optimiziraju različite kriterije [103].

#### 2.4.3 PSO za konstrukciju grafa Bajesove mreže iz podataka

PSO je u nekoliko aktualnih publikacija [91][120][121][122][123][91][124][125][126] prikazan kao meta-heuristika koja je izuzetno pogodna za primjenu u metričkim algoritmima za učenje strukture Bajesove mreže. Sve navedene publikacije se zasnivaju na istoj ideji – upotrebom PSO meta-heuristike pronaći onu strukturu koja maksimizira neku funkciju sličnosti strukture BN-a i podataka na osnovu kojih se konstruiše graf. Razlike se pronalaze u samoj implementaciji PSO algoritma ili neke njene modifikacije. Preciznije rečeno, razlike su uglavnom u pristupu reprezentacije potencijalnih rješenja i mehanizmima zadovoljavanja ograničenja tj. osiguravanja ispravnosti grafa. Mehanizmi zaustavljanja se zasnivaju na broju iteracija ili stepenu poboljšanja rješenja. Korištene ciljne funkcije uglavnom pripadaju nekom od ranije navedenih tipova funkcija, pri čemu je najčešće korištena Bajesova distribucija vjerovatnoće. Iako autori ovih članaka tvrde da ovi algoritmi pokazuju odlične rezultate, ipak treba imati na umu da su isti uglavnom primijenjeni na *benchmark* problemima, a ne na kompleksnim problemima iz realnog sektora.

J. Cowie i saradnici [121] su uveli dva pristupa CONAR i REST. U oba slučaja graf prikazuju matricom povezanosti čvorova a ciljna funkcija je preuzeta iz K2 [62] algoritma. Razlika u ova dva pristupa je samo u mehanizmima za osiguravanje validnosti grafa. Dok se prvi algoritam oslanja na slučajni izbor grafa i metodom prepravljanja grafa osigurava zadovoljavanje ograničenja, drugi pristup koristi metode za generisanje validnih grafova već u fazi izbora tačke, tj. u fazi kreiranja matrice povezanosti. Ove dvije implementacije ne pokazuju velike razlike u performansama kada je riječ o primjeni na *benchmark* primjerima.

M. Valian i saradnici [124] koriste PSO koji također koristi matricu povezanosti za reprezentaciju potencijalnih rješenja i mehanizam odbacivanja rješenja ukoliko novonastala matrica ne ispunjava sva ograničenja validnosti grafa. Ovaj algoritam je uspješno primijenjen na jedan od često korištenih *benchmark* problema, ASIA.

W. Chun-Feng i saradnici u svojoj publikaciji [125] predstavljaju graf u vidu matrice povezanosti i koriste funkciju cilja baziranu na Bajesovoj raspodjeli vjerovatnoće. Specifičnost ovog pristupa leži u heuristici koja je korištena. Naime, ovaj algoritam je hibrid optimizacije rojem čestica i kolonije vještačkih pčela (engl. *Artificial Bee Colony, ABC*). Ovaj hibridni algoritam koristi PSO formulu za računanje brzine čestice, a topologiju i implementaciju grupne inteligencije ABC-a.

N. Forthier i saradnici u [126] analiziraju kako paralelni PSO može biti korišten za učenje grafa. Ovaj pristup koristi odvojene populacije za pronalazak pod-strukture za svaku varijablu u podacima, te potom vrši povezivanje dobivenih pod-grafova. Za razliku od ostalih istraživanja, ovdje je korištena i BIC (engl. *Bayesian Information Criterion*) [92] mjera kao ciljna funkcija.

### 3 Motivacija za istraživanje

Gotovo sve oblasti u kojima je potrebno upravljati kompleksnim sistemima, susreću se s problemom neoptimalnog donošenja odluka. Nepreglednost problemskog domena je jedan od najvećih problema jer dovodi do toga da donosioci odluka nisu u stanju predvidjeti efekte koje će njihove odluke imati na

sistem. Prilikom provođenja inicijalnih istraživanja u sklopu ove disertacije uočeno je da se sa ovim problemom susreću i eksperti iz oblasti upravljanja prirodnim resursima, te da je za rješavanje date problematike potreban alat za modeliranje problemskog domena koji se može koristiti za simulaciju akcija i predikciju vjerovatnih posljedica. Iz ove potrebe proizlazi motivacija da se u sklopu ove disertacije istraži oblast Bajesovih mreža te da se iste iskoriste u svrhu optimalnijeg upravljanja unutar navedenog domena.

U ovoj disertaciji će u kooperaciji s ekspertima iz područja ekologije i genetike biti analiziran problem iz oblasti upravljanja prirodnim resursima. Bajesove mreže su prepoznate kao prostor za istraživanje koji, imajući u vidu sve prethodno rečeno, djeluje obećavajuće u smislu pronalaska kvalitetnog rješenja za datu problematiku. Dati problem zasnovan je na činjenici da opstanak i razvoj neke vrste ovisi o mnogim faktorima, te da za kvalitetno donošenje odluka u datoj oblasti upravitelj mora imati uvid u interakciju datih faktora i njihov uticaj na domen. Ovako kompleksni sistemi su podložni naglim promjenama i imaju visok stepen neizvjesnosti. S obzirom da nije moguće definisati preslikavanja s kojima se može opisati domen i klasičnim metodama pronaći optimalne parametre sistema, uočen je potencijal kauzalnih struktura za vizualizaciju i simulaciju.

U saradnji sa ekspertima iz oblasti genetike uočeno je da se odluke koje se donose pri upravljanju prirodnim resursima uglavnom zasnivaju na znanju, iskustvu i intuiciji pojedinca koji upravlja prirodnim resursima. U praksi se uglavnom ne primjenjuje nikakva metoda koja bi objektivno podržala ili osporila ekspertnu procjenu ili predloženu akciju. Različiti eksperti mogu davati različite prijedloge za rješenje problema, a efekti poduzetih akcija se nekad primjećuju tek nakon dužeg vremenskog perioda. Iz ovog razloga je izražena potreba za alatom koji na osnovu historijskih podataka o poduzetim akcijama i ishodima, te podacima koji opisuju relevantne aspekte domena, može simulirati događaje u datom domenu. Alat je zasnovan na podacima vođenom algoritmu koji na osnovu definisane baze podataka može formirati model domena. S obzirom da je proces kreiranja modela univerzalan, izmjenom baze podataka moguće je na isti način kreirati model drugog domena. Kako bi se dati model mogao koristiti za vizualizaciju, simulaciju i predikciju vjerovatnih događaja, izabran je pristup kreiranja modela u obliku grafa Bajesove mreže. Tokom pregleda literature uočeno je nekoliko publikacija koje se fokusiraju na upotrebu Bajesovih mreža za rješavanje sličnih problema iz pomenute oblasti [34]-[42]. Konkretno, analiza vodenih staništa i razvoj populacija riba u istim, te uticaj neselektivnog podmlađivanja i izgradnje brana na posmatranom toku je višestruko obrađivana tema. Ova disertacija će se također fokusirati na analizu procesa i uticaja genetskih, ekoloških i socioekonomskih faktora na konkretni vodeni tok i posmatranu vrstu.

Da bi se kreirala Bajesova mreža za dati domen prvo je potrebno prikupiti podatke koji opisuju sistem a zatim identificirati model iz istih. Definisanje i prikupljanje grupa podataka koji se mogu koristiti za konstrukciju modela je zadatak koji je u nadležnosti domenskog eksperta. On poznaje problemski domen i procese koji se odvijaju u istom, te može odrediti relevantne izvore podataka. Međutim, proces konstrukcije modela sadrži komponentu koja se može predstaviti kao optimizacijski problem, te je stoga potrebno i ekspertno znanje i iz oblasti optimizacije. U ovoj disertaciji će upravo aspekt definisanja efektivnih metoda za identifikaciju strukture grafa Bajesove mreže iz podataka biti glavni zadatak autora. Kada se definiše struktura grafa i parametri modela, moguće je trenirati model i koristiti ga za simulaciju i predikciju.

Tokom pregleda literature primijećeno je da postoji široka paleta algoritama koji nastoje identificirati strukturu Bajesove mreže, te su se neki od njih pokazali kao efikasni za rješavanje jednostavnih problema. Međutim, uočeno je i dosta prostora za dodatna istraživanja i poboljšanja, posebno u pogledu generisanja

strukture grafa za kompleksne domene. Iako postoje mnoga otvorena pitanja u ovoj oblasti, u ovoj disertaciji će fokus biti na pronalasku odgovora na neka od sljedećih pitanja.

- Da li se kompleksni sistemi kao prirodna staništa sa više vrsta mogu kvalitetno modelirati Bajesovom mrežom? Da li takva Bajesova mreža može u razumnom vremenu vršiti pouzdane simulacije i predikcije događaja u definisanom domenu.
- Da li postoji jedinstveno okruženje koje povezuje bazu domenskog znanja sa algoritmom za generisanje grafa iz podataka? Naime, iako već postoje neka softverska rješenja za rad s Bajesovim mrežama, većina njih su fokusirana na procese interferencije, a ne samog generisanja strukture grafa. Okruženja koja pak generišu graf su s druge strane često efikasna samo za manje probleme i uglavnom nemaju mogućnost uključivanja preferencija eksperta u realnom vremenu.
- Da li je moguće korištenjem domenskih podataka i preferencija upravitelja generisati graf za Bajesove mreže za odlučivanje, DBN? Da li je moguće u samom procesu kreiranja strukture uključiti čvorove za prikazivanje funkcije korisnosti na osnovu koje se u budućnosti mogu donositi odluke?
- Da li je moguće koristiti PSO meta-heuristiku za razvoj efikasnog metričkog algoritma za vrlo kompleksne domene?
- Da li je moguće unaprijediti mehanizme PSO algoritma za efikasnije pretraživanje diskretnog ograničenog prostora? Prema aktualnoj literaturi, za reprezentaciju potencijalnih struktura se uglavnom koristi matrica povezanosti čvorova. Da li bi strukture koje bi predstavljale svojevrsne rječičke roditeljskih čvorova postizale bolje performanse? Da li mehanizmi pomaka koji vode računa o validnosti strukture daju bolje rezultate od pristupa odbacivanja nevalidnih struktura?
- Da li je moguće unaprijediti mehanizam zaustavljanja PSO-a korištenjem neke od mjera efikasnosti za isti?
- Koja ciljna funkcija može najbolje opisati podudarnost grafa i podataka? Da li uključivanjem dva odvojena kriterija i korištenjem MOPSO algoritma možemo unaprijediti performanse metričkih metoda? Da li bi korištenje kriterija minimizacije grafa pored kriterija podudarnosti podataka rezultiralo poboljšanim grafom?
- Da li je moguće koristiti hibrid paralelnog i pareto MOPSO-a za rješavanje ovakvog višekriterijalnog problema?
- Koja topologija PSO-a daje najbolje performanse za ovaj problem?

## 4 Ciljevi i plan istraživanja

Primarni cilj doktorske disertacije proizlazi iz pregleda stanja u oblasti istraživanja, lične motivacije i analize realnog problema iz oblasti upravljanja prirodnim resursima. Cilj koji se nameće za ovu disertaciju je da se definisanjem varijante PSO algoritma ili neke njegove modifikacije, kreira efektivan podacima vođen algoritam za generisanje grafa Bajesove mreže.

Istraživanje će biti sprovedeno prema sljedećem planu:

1. Analiza stanja u oblasti istraživanja i izrada detaljnog pregleda literature uz kritički osvrt na postojeće pristupe, njihove prednosti i nedostatke, te ostvarene rezultate na polju prepoznavanja strukture Bajesove mreže iz podataka. Poseban značaj će biti dat metričkim metodama s fokusom na one koje koriste PSO meta-heuristiku.

2. Formiranje simulacionog okruženja koje će omogućiti implementaciju i validaciju PSO baziranog, podacima vođenog algoritma za detekciju strukture BN-a.
3. Sistematično ispitivanje uticaja korištenja različitih varijanti nekih strukturnih elemenata PSO-a na performanse algoritma. Poseban pažnja će biti data mehanizmima za reprezentaciju potencijalnih rješenja i zadovoljenja ograničenja.
4. Sistematično ispitivanje i validiranje uticaja izabrane ciljne funkcije na performanse datog algoritma i tačnost generisanog grafa.
5. Sistematično ispitivanje mogućnosti upotrebe neke od mjera efikasnosti metričkih algoritama kao mehanizma zaustavljanja u datom algoritmu.
6. Analiza uticaja primjene MOPSO koncepta na performanse datog algoritma.
7. Priprema baze podataka koji će se koristiti za modeliranje realnog problema iz oblasti upravljanja prirodnim resursima.
8. Kreiranje modela posmatranog realnog domena korištenjem pomenutog algoritma.
9. Formiranje okruženja za rad s identificiranom Bajesovom mrežom koje će biti korišteno za simulaciju i predikciju vjerovatnih događaja u problemskom domenu.

## 5 Metodologija istraživanja

Kako bi se mogli postići navedeni ciljevi i izvršiti pometnuta istraživanja, neophodno je kreirati okruženje za razvoj i ispitivanje. S obzirom da ova disertacija obrađuje Bajesove mreže s dva različita aspekta, istraživanje će biti vršeno kroz dvije etape korištenjem dva odvojena razvojna okruženja. Prva etapa će se fokusirati na razvoj algoritma za generisanje strukture Bajesove mreže, dok će u drugoj etapi fokus biti na korištenju definisane strukture Bajesove mreže za predikciju i simulaciju.

### 5.1 Okruženje za razvoj i poređenje metričkih algoritama za detekciju modela

Za implementaciju metričkog algoritma te pohranjivanje podataka o domenu i bilješki o performansama biti će korišteno okruženje Visual Studio 2015 [127] sa C# programskim jezikom i bazom podataka SQL Server 2014 [128]. Validacija i mjerenje performansi algoritma će biti rađeno prvenstveno na *benchmark* primjerima iz ove oblasti. U ovoj disertaciji će biti korišteni *benchmark* primjeri objavljeni u [129] jer su korišteni u mnogim istraživanjima i komparacijama metričkih algoritama. Velika prednost ove kolekcije je što sadrži primjere različite složenosti, od onih jednostavnih kao *ASIA*, *Cancer* dr., sa desetak čvorova i grana u grafu, do složeniji primjera kao *Andes*, *Diabetes*, *Pathfinder* i dr., koji opisuju kompleksne domene sa čak više stotina čvorova i grana koji modeliraju domen. Navedeni izvor za sve odabrane primjere sadrži opis, parametre mreže te graf mreže za koji se smatra da najbolje opisuje domen. Zato se za ove primjere može reći da je ispravno rješenje problema poznato, te se rezultati algoritama za generisanje strukture grafa mogu validirati tako što se dobiveni graf poredi sa ranije definisanim ispravnim grafom. Za poređenje grafova će biti korištene neke od mjera performansi navedene u poglavlju 2.3.2.

Da bi se neki metrički algoritam mogao primijeniti na nekom od *benchmark* problema, neophodno je imati bazu podataka koji opisuju domen problema. S obzirom da [129] ne nudi ovakve podatke, isti će za date *benchmark* probleme biti preuzeti iz softverskog okruženja *R* [130][131]. Ovi podaci će biti uneseni u testno okruženje i korišteni za kreiranje odgovarajućeg grafa Bajesove mreže. Validacijom algoritma na ovakvim primjerima će biti moguće ustanoviti nedostatke istog te vršiti stalne dorade dok se ne postignu željeni rezultati. Finalna verzija datog algoritma će potom biti primijenjena na realnom problemu.

Realni problem iz oblasti upravljanja prirodnim resursima koji se razmatra u ovoj disertaciji će najvjerovatnije biti temeljen na prikupljenim podacima vezanim za populaciju ribe pastrmke u gornjem toku rijeke Neretve. Baza podataka koja će biti formirana za ovaj domen treba da sadrži podatke koji predstavljaju tri aspekta domena – genetski, ekološki i socioekonomski. Na ovaj način će se uzeti u obzir genetska predispozicija vrste, uticaj ekoloških faktora kao pojava druge vrste u datom toku, te uticaj izgradnje brana i energetskih sistema na prirodne resurse u ovom području. Relevantni izvori podataka će biti identificirani od strane domenskog eksperta, te će podaci nakon primjene eventualnih transformacijskih procesa za čišćenje i dopunu biti uneseni u opisano testno okruženje. Implementirani algoritam će se onda primijeniti na ove podatke u svrhu generisanja Bajesove mreže za dati domen.

## 5.2 Okruženje za rad sa Bajesovom mrežom

Nakon što se za navedeni realni problem generiše model u obliku Bajesove mreže, isti je potrebno prikazati u okruženju za rad s Bajesovim mrežama. Postoje mnogi softverski paketi koji omogućavaju vizualno prikazivanje, konfigurisanje i treniranje mreže. U ovoj disertaciji će konkretan softverski alat za rad sa Bajesovom mrežom biti odabran u kasnijoj fazi izrade zadatka. Isti će biti odabran u saradnji sa domenskim ekspertnim timom u skladu s njihovim potrebama i zahtjevima koji nameće kompleksnost definisanog problema.

Izabrani program će se prvo koristiti za vizualizaciju modela te će se vršiti konfiguracija tabela uslovne vjerovatnoće. Algoritam za interferenciju će biti odabran iz skupa dostupnih algoritama u datom okruženju, a treniranje će se vršiti sa postojećim testnim skupom historijskih podataka. Kad se Bajesova mreža istrenira tako da rezultati budu zadovoljavajući, ista će se početi primjenjivati za rad s novim podacima i dokazima. Data mreža će se moći koristiti za simuliranje i predikciju vjerovatnih događaja što će pomoći donosiocima odluka da odaberu optimalne akcije.

## 6 Očekivani izvorni naučni doprinos disertacije

Nakon analize postojećeg stanja u oblasti i rezultata preliminarne istraživanja, identificirani su sljedeći potencijalni naučni doprinosi disertacije:

- Poboljšanje postojećeg ili formiranje novog, podacima vođenog, PSO baziranog algoritma za detekciju strukture grafa Bajesove mreže, u cilju poboljšavanja performansi i tačnosti automatiziranog procesa kreiranja modela problemskog domena.
- Aplikacija datog algoritma na realni problem iz sektora upravljanja prirodnim resursima u svrhu ispitivanja mogućnosti primjene Bajesove mreže za kvalitetnije donošenje odluka.

Pored navedenih doprinosa, disertacija će ponuditi sistematičan pregled literature te analizirati prednosti i nedostatke postojećih formalno definisanih koncepata unutar oblasti. Data analiza će zasigurno rezultirati s otvaranjem novih pitanja na koja sama disertacija neće moći ponuditi odgovore. Ipak, samo postavljanje datih pitanja može poslužiti kao poticaj za dalja istraživanja u oblasti Bajesovih mreža, što se također može smatrati doprinosom.

## 7 Polazna literatura

- [1] Grover, Jeff. "A Literature Review of Bayes' Theorem and Bayesian Belief Networks (BBN)." In *Strategic Economic Decision-Making*, pp. 11-27. Springer New York, 2013.
- [2] Korb, Kevin B., and Ann E. Nicholson. *Bayesian artificial intelligence*. CRC press, 2010.
- [3] Margaritis, Dimitris. "Learning Bayesian network model structure from data." PhD diss., US Army, 2003.
- [4] Pellet, Jean-Philippe, and André Elisseeff. "Using Markov blankets for causal structure learning." *Journal of Machine Learning Research* 9, no. Jul (2008): 1295-1342.
- [5] Geiger, Dan, Tom S. Verma, and Judea Pearl. "d-separation: From theorems to algorithms." *arXiv preprint arXiv:1304.1505* (2013).
- [6] Frydenberg, Morten. "The chain graph Markov property." *Scandinavian Journal of Statistics* (1990): 333-353.
- [7] Olesen, Kristian G., Uffe Kjaerulff, Frank Jensen, Finn V. Jensen, Bjoern Falck, Steen Andreassen, and Stig K. Andersen. "A multinomial network for the median nerve—a case study on loops." *Applied Artificial Intelligence an International Journal* 3, no. 2-3 (1989): 385-403.
- [8] Heckerman, David E., Eric J. Horvitz, and Bharat N. Nathwani. "Toward normative expert systems: The Pathfinder project." *Methods of information in medicine* 31 (1991): 90-105.
- [9] Kahn, Charles E., Linda M. Roberts, Katherine A. Shaffer, and Peter Haddawy. "Construction of a Bayesian network for mammographic diagnosis of breast cancer." *Computers in biology and medicine* 27, no. 1 (1997): 19-29.
- [10] Onisko, Agnieszka, Marek J. Druzdzel, and Hanna Wasyluk. "A probabilistic causal model for diagnosis of liver disorders." In *Proceedings of the Seventh International Symposium on Intelligent Information Systems (IIS-98)*, pp. 379-387. 1998.
- [11] van der Gaag, Linda C., Silja Renooij, C. L. M. Witteman, Berthe MP Aleman, and Babs G. Taal. "Probabilities for a probabilistic network: a case study in oesophageal cancer." *Artificial Intelligence in medicine* 25, no. 2 (2002): 123-148.
- [12] Burnside, Elizabeth S., Edward A. Sickles, Rita E. Sohlich, and Katherine E. Dee. "Differential value of comparison with previous examinations in diagnostic versus screening mammography." *American Journal of Roentgenology* 179, no. 5 (2002): 1173-1177.
- [13] Charitos, Theodore, Linda C. Van Der Gaag, Stefan Visscher, Karin AM Schurink, and Peter JF Lucas. "A dynamic Bayesian network for diagnosing ventilator-associated pneumonia in ICU patients." *Expert Systems with Applications* 36, no. 2 (2009): 1249-1258.
- [14] Cowell, Robert G., A. Philip Dawid, T. Hutchinson, and David J. Spiegelhalter. "A Bayesian expert system for the analysis of an adverse drug reaction." *Artificial Intelligence in Medicine* 3, no. 5 (1991): 257-270.
- [15] Kline, Jeffrey A., Andrew J. Novobilski, Christopher Kabrhel, Peter B. Richman, and D. Mark Courtney. "Derivation and validation of a Bayesian network to predict pretest probability of venous thromboembolism." *Annals of emergency medicine* 45, no. 3 (2005): 282-290.
- [16] Štajduhar, Ivan, Bojana Dalbelo-Bašić, and Nikola Bogunović. "Impact of censoring on learning Bayesian networks in survival modelling." *Artificial intelligence in medicine* 47, no. 3 (2009): 199-217.
- [17] Smith, Carl S., Alison L. Howes, Bronwyn Price, and Clive A. McAlpine. "Using a Bayesian belief network to predict suitable habitat of an endangered mammal—The Julia Creek dunnart (*Sminthopsis douglesi*)." *Biological Conservation* 139, no. 3 (2007): 333-347.
- [18] Aitkenhead, M. J., and I. H. Aalders. "Predicting land cover using GIS, Bayesian and evolutionary algorithm methods." *Journal of environmental Management* 90, no. 1 (2009): 236-250.

- [19]Zhang, Harry. "The optimality of naive Bayes." AA 1, no. 2 (2004): 3.
- [20]Bielza, Concha, and Pedro Larrañaga. "Discrete Bayesian network classifiers: a survey." ACM Computing Surveys (CSUR) 47, no. 1 (2014): 5.
- [21]Friedman, Nir, Dan Geiger, and Moises Goldszmidt. "Bayesian network classifiers." Machine learning 29, no. 2-3 (1997): 131-163.
- [22]Guo, Haipeng, and William Hsu. "A survey of algorithms for real-time Bayesian network inference." In AAAI/KDD/UAI02 Joint Workshop on Real-Time Decision Support and Diagnosis Systems. Edmonton, Canada, 2002.
- [23]Liu, Yi, Nan Qiao, Shanshan Zhu, Ming Su, Na Sun, Jerome Boyd-Kirkup, and Jing-Dong J. Han. "A novel Bayesian network inference algorithm for integrative analysis of heterogeneous deep sequencing data." Cell research 23, no. 3 (2013): 440-443.
- [24]Marbach, Daniel, James C. Costello, Robert Küffner, Nicole M. Vega, Robert J. Prill, Diogo M. Camacho, Kyle R. Allison et al. "Wisdom of crowds for robust gene network inference." Nature methods 9, no. 8 (2012): 796-804.
- [25]Larrañaga, Pedro, Hossein Karshenas, Concha Bielza, and Roberto Santana. "A review on evolutionary algorithms in Bayesian network learning and inference tasks." Information Sciences 233 (2013): 109-125.
- [26]Shachter, Ross D., and Mark Alan Peot. "Simulation approaches to general probabilistic inference on belief networks." arXiv preprint arXiv:1304.1526 (2013).
- [27]Gogate, Vibhav, and Rina Dechter. "Approximate inference algorithms for hybrid bayesian networks with discrete constraints." arXiv preprint arXiv:1207.1385 (2012).
- [28]Taroni, Franco, Alex Biedermann, Silvia Bozza, Paolo Garbolino, and Colin Aitken. "Bayesian decision networks." Bayesian Networks for Probabilistic Inference and Decision Analysis in Forensic Science (2014): 343-369.
- [29]Zhu, J. Y., and A. Deshmukh. "Application of Bayesian decision networks to life cycle engineering in Green design and manufacturing." Engineering Applications of Artificial Intelligence 16, no. 2 (2003): 91-103.
- [30]Murphy, Kevin Patrick. "Dynamic bayesian networks: representation, inference and learning." PhD diss., University of California, Berkeley, 2002.
- [31]Murphy, Kevin P. "Dynamic bayesian networks." Probabilistic Graphical Models, M. Jordan 7 (2002).
- [32]Kasper, Dietmar, GaliaWeidl, Thao Dang, Gabi Breuel, Andreas Tamke, Andreas Wedel, and Wolfgang Rosenstiel. "Object-oriented Bayesian networks for detection of lane change maneuvers." IEEE Intelligent Transportation Systems Magazine 4, no. 3 (2012): 19-31.
- [33]Langseth, Helge, and Olav Bangsø. "Parameter learning in object-oriented Bayesian networks." Annals of Mathematics and Artificial Intelligence 32, no. 1-4 (2001): 221-243.
- [34]Pollino, Carmel A., Owen Woodberry, Ann Nicholson, Kevin Korb, and Barry T. Hart. "Parameterisation and evaluation of a Bayesian network for use in an ecological risk assessment." Environmental Modelling & Software 22, no. 8 (2007): 1140-1152.
- [35]Pollino, C. A., and C. Henderson. "Bayesian networks: A guide for their application in natural resource management and policy." Landscape Logic, Technical Report 14 (2010).
- [36]Hart, Barry T., P. S. Lake, J. Angus Webb, and Michael R. Grace. "Ecological risk to aquatic systems from salinity increases." Australian Journal of Botany 51, no. 6 (2003): 689-702.
- [37]Bromley, John, Nick A. Jackson, O. J. Clymer, Anna Maria Giacomello, and Finn Verner Jensen. "The use of Hugin® to develop Bayesian networks as an aid to integrated water resource planning." Environmental Modelling & Software 20, no. 2 (2005): 231-242.

- [38]Jensen, F. V., U. Kjærulff, K. G. Olesen, and J. Pedersen. Et forprojektil et ekspertsystem for drift af spildevandsrensning (a prototype expert system for control of waste water treatment). Technical report, JudexDatasytemer A/S, Aalborg, Denmark, 1989. In Danish, 1989.
- [39]Varis, Olli, and Marko Keskinen. "Policy analysis for the Tonle Sap Lake, Cambodia: a Bayesian network model approach." *Water Resources Development* 22, no. 3 (2006): 417-431.
- [40]Varis, Olli, and VirpiLahtela. "Integrated water resources management along the Senegal River: introducing an analytical framework." *International Journal of Water Resources Development* 18, no. 4 (2002): 501-521.
- [41]Barton, D. N., T. Saloranta, S. J. Moe, H. O. Eggestad, and S. Kuikka. "Bayesian belief networks as a meta-modelling tool in integrated river basin management—Pros and cons in evaluating nutrient abatement decisions under uncertainty in a Norwegian river basin." *Ecological Economics* 66, no. 1 (2008): 91-104.
- [42]McCann, Robert K., Bruce G. Marcot, and Rick Ellis. "Bayesian belief networks: applications in ecology and natural resource management." *Canadian Journal of Forest Research* 36, no. 12 (2006): 3053-3062.
- [43]Fenton, Norman E., and Martin Neil. "Decision support software for probabilistic risk assessment using Bayesian networks." (2014).
- [44]Conrady, Stefan, and Lionel Jouffe. "Introduction to Bayesian Networks & BayesiaLab." September 3, no. 201 (2013): 3.
- [45]Murphy, Kevin. "Bayesian Network Toolbox (BNT),(Version 1.0. 7) Computer software." (2014).
- [46]Pérez-Ariza, Cora Beatriz, Ann E. Nicholson, Kevin B. Korb, Steven Mascaro, and Chao Heng Hu. "Causal discovery of dynamic Bayesian networks." In *Australasian Joint Conference on Artificial Intelligence*, pp. 902-913. Springer Berlin Heidelberg, 2012.
- [47]Srinivas, Sampath, and John S. Breese. "IDEAL: A software package for analysis of influence diagrams." *arXiv preprint arXiv:1304.1107* (2013).
- [48]Jing, Chen, Fu Jingqi, and Su Wei. "Model Construction and Function Realizing of Bayesian Network Based on Netica Platform." *International Journal of Advancements in Computing Technology* 5, no. 6 (2013).
- [49]Druzdzel, Marek J. "Genie Smile,(Version 20.)[Computer software]." (2014).
- [50]Glymour, C., R. Scheines, P. Spirtes, and J. Ramsey. "TETRAD project." (2012).
- [51]Oboler, Andre. The kebn process: A new approach to knowledge engineering with bayesian nets. Technical report, Monash University, 2002.
- [52]Cao, Yonghui. "Study of Four Types of Learning Bayesian Networks Cases." *Appl. Math* 8, no. 1 (2014): 379-386.
- [53]Spiegelhalter, David J., and Steffen L. Lauritzen. "Sequential updating of conditional probabilities on directed graphical structures." *Networks* 20, no. 5 (1990): 579-605.
- [54]Almond, Russell G. "An irt-based parameterization for conditional probability tables." In *Bayesian Modelling Application Workshop at the Uncertainty in Artificial Intelligence (UAI) Conference*, Amsterdam, The Netherlands. 2015.
- [55]de Morais Andrade, Pablo, Julio Michael Stern, and Carlos Alberto de Bragança Pereira. "Bayesian test of significance for conditional independence: the multinomial model." *Entropy* 16, no. 3 (2014): 1376-1395.
- [56]Vomlel, Jiri. "A generalization of the noisy-or model to multivalued parent variables." In *Proc. 16th Czech-Japan Seminar on Data Analysis and Decision Making under Uncertainty*, pp. 19-27. 2013.
- [57]Burge, John, and Terran Lane. "Shrinkage estimator for Bayesian network parameters." In *European Conference on Machine Learning*, pp. 67-78. Springer Berlin Heidelberg, 2007.

- [58] O'Donnell, Rodney T., Lloyd Allison, and Kevin B. Korb. "Learning hybrid Bayesian networks by MML." In Australasian Joint Conference on Artificial Intelligence, pp. 192-203. Springer Berlin Heidelberg, 2006.
- [59] Marlin, Benjamin M., and Richard S. Zemel. "Collaborative prediction and ranking with non-random missing data." In Proceedings of the third ACM conference on Recommender systems, pp. 5-12. ACM, 2009.
- [60] Sebastiani, Paola, and Marco Ramoni. "Bayesian inference with missing data using bound and collapse." *Journal of Computational and Graphical Statistics* 9, no. 4 (2000): 779-800.
- [61] Friedman, Nir. "The Bayesian structural EM algorithm." In Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence, pp. 129-138. Morgan Kaufmann Publishers Inc., 1998.
- [62] Acid, Silvia, Luis M. de Campos, Juan M. Fernández-Luna, Susana Rodríguez, José María Rodríguez, and José Luis Salcedo. "A comparison of learning algorithms for Bayesian networks: a case study based on data from an emergency medical service." *Artificial intelligence in medicine* 30, no. 3 (2004): 215-232.
- [63] Xiong, Wei, Yonghui Cao, and Hui Liu. "Study of Bayesian Network Structure Learning." *Appl. Math* 7, no. 1L (2013): 49-54.
- [64] Malone, Brandon, Matti Järvisalo, and Petri Myllymäki. "Impact of learning strategies on the quality of Bayesian networks: An empirical evaluation." In Proceedings of the 31st Conference on Uncertainty in Artificial Intelligence (UAI 2015). 2015.
- [65] Spirtes, P., C. Glymour, and R. Scheines. "Lecture Notes in Statistics." Causation, prediction, and search 81 (1993).
- [66] Cheng, Jie, David A. Bell, and Weiru Liu. "An algorithm for Bayesian belief network construction from data." In proceedings of AI & STAT'97, pp. 83-90. 1997.
- [67] De Campos, Luis M., and Juan F. Huete. "A new approach for learning belief networks using independence criteria." *International Journal of Approximate Reasoning* 24, no. 1 (2000): 11-37.
- [68] Coello, Carlos Coello, Gary B. Lamont, and David A. Van Veldhuizen. *Evolutionary algorithms for solving multi-objective problems*. Springer Science & Business Media, 2007.
- [69] Marler, R. Timothy, and Jasbir S. Arora. "Survey of multi-objective optimization methods for engineering." *Structural and multidisciplinary optimization* 26, no. 6 (2004): 369-395.
- [70] Heckerman, David, Dan Geiger, and David M. Chickering. "Learning Bayesian networks: The combination of knowledge and statistical data." *Machine learning* 20, no. 3 (1995): 197-243.
- [71] Cooper, Gregory F., and Edward Herskovits. "A Bayesian method for the induction of probabilistic networks from data." *Machine learning* 9, no. 4 (1992): 309-347.
- [72] Herskovits, Edward H., and Gregory F. Cooper. "Kutato: An entropy-driven system for construction of probabilistic expert systems from databases." *arXiv preprint arXiv:1304.1088* (2013).
- [73] Bouckaert, Remco R. "Probabilistic network construction using the minimum description length principle." In European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty, pp. 41-48. Springer Berlin Heidelberg, 1993.
- [74] Lam, Wai, and Fahiem Bacchus. "Learning Bayesian belief networks: An approach based on the MDL principle." *Computational intelligence* 10, no. 3 (1994): 269-293.
- [75] O'donnell, R. T., A. E. Nicholson, B. Han, K. B. Korb, M. J. Alam, and L. R. Hope. "Incorporating expert elicited structural information in the CaMML causal discovery program." In Proceedings of the 19th Australian Joint Conference on Artificial Intelligence: Advances in Artificial Intelligence, pp. 1-16. 2006.
- [76] DALY, RÓNÁN, QIANG SHEN, STUART AITKEN, TANGMING YUAN, DAVID MOORE, CHRIS REED, ANDREW RAVENSCROFT et al. "The knowledge engineering."

- [77]Cooper, Gregory F., and Changwon Yoo. "Causal discovery from a mixture of experimental and observational data." In Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence, pp. 116-125. Morgan Kaufmann Publishers Inc., 1999.
- [78]Acid, S., and L. M. de Campos. "An algorithm for learning probabilistic belief networks using minimum dseparating sets." Inf. Téc. DECSAI-00001, Dpto de Ciencias de la Computación– Universidad de Granada (2001).
- [79]Friedman, Nir, IftachNachman, and Dana Peér. "Learning bayesian network structure from massive datasets: the «sparse candidate «algorithm." In Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence, pp. 206-215. Morgan Kaufmann Publishers Inc., 1999.
- [80]Teyssier, Marc, and Daphne Koller. "Ordering-based search: A simple and effective algorithm for learning Bayesian networks." arXiv preprint arXiv:1207.1429 (2012).
- [81]Teyssier, Marc, and Daphne Koller. "Ordering-based search: A simple and effective algorithm for learning Bayesian networks." arXiv preprint arXiv:1207.1429 (2012).
- [82]Malone, Brandon, and Changhe Yuan. "Evaluating anytime algorithms for learning optimal Bayesian networks." arXiv preprint arXiv:1309.6844 (2013).
- [83]Zhang, Yinghua, Wensheng Zhang, and Yuan Xie. "Improved heuristic equivalent search algorithm based on maximal information coefficient for Bayesian network structure learning." Neurocomputing 117 (2013): 186-195.
- [84]Madsen, Anders L., Frank Jensen, Antonio Salmerón, Helge Langseth, and Thomas D. Nielsen. "A parallel algorithm for Bayesian network structure learning from large data sets." Knowledge-Based Systems (2016).
- [85]De Campos, Luis M., Juan M. Fernandez-Luna, José A. Gámez, and José M. Puerta. "Ant colony optimization for learning Bayesian networks." International Journal of Approximate Reasoning 31, no. 3 (2002): 291-311.
- [86]Larrañaga, Pedro, Mikel Poza, YosuYurramendi, Roberto H. Murga, and Cindy M. H. Kuijpers. "Structure learning of Bayesian networks by genetic algorithms: A performance analysis of control parameters." IEEE transactions on pattern analysis and machine intelligence 18, no. 9 (1996): 912-926.
- [87]Özçift, Akın, and ArifGülten. "Genetic algorithm wrapped Bayesian network feature selection applied to differential diagnosis of erythemato-squamous diseases." Digital Signal Processing 23, no. 1 (2013): 230-237.
- [88]Pelikan, Martin, David E. Goldberg, and Erick Cantú-Paz. "BOA: The Bayesian optimization algorithm." In Proceedings of the 1st Annual Conference on Genetic and Evolutionary Computation-Volume 1, pp. 525-532. Morgan Kaufmann Publishers Inc., 1999.
- [89]Campos, Luis M. de. "Independency relationships and learning algorithms for singly connected networks." Journal of Experimental & Theoretical Artificial Intelligence 10, no. 4 (1998): 511-549.
- [90]Schwarz, Gideon. "Estimating the dimension of a model." The annals of statistics 6, no. 2 (1978): 461-464.
- [91]Correa, Elon S., Alex A. Freitas, and Colin G. Johnson. "Particle swarm and bayesian networks applied to attribute selection for protein functional classification." In Proceedings of the 9th annual conference companion on Genetic and evolutionary computation, pp. 2651-2658. ACM, 2007.
- [92]O'Donnell, Rodney T. Flexible Causal Discovery with MML. Monash University, 2010.
- [93]Kim, Young-Gyun, and Marco Valtorta. "On the detection of conflicts in diagnostic Bayesian networks using abstraction." In Proceedings of the Eleventh conference on Uncertainty in artificial intelligence, pp. 362-367. Morgan Kaufmann Publishers Inc., 1995.

- [94]Eberhart, Russ C., and James Kennedy. "A new optimizer using particle swarm theory." In Proceedings of the sixth international symposium on micro machine and human science, vol. 1, pp. 39-43. 1995.
- [95]Clerc, M. "Standard particle swarm optimisation from 2006 to 2011." Particle Swarm Central 253 (2011).
- [96]Zambrano-Bigiarini, Mauricio, Maurice Clerc, and Rodrigo Rojas. "Standard particle swarm optimization 2011 at cec-2013: A baseline for future pso improvements." In 2013 IEEE Congress on Evolutionary Computation, pp. 2337-2344. IEEE, 2013.
- [97]Clerc, Maurice. "Standard particle swarm optimisation." (2012).
- [98]Clerc, Maurice, and James Kennedy. "The particle swarm-explosion, stability, and convergence in a multidimensional complex space." IEEE transactions on Evolutionary Computation 6, no. 1 (2002): 58-73.
- [99]Niknam, T., M. R. Narimani, J. Aghaei, and R. Azizipanah-Abarghooee. "Improved particle swarm optimisation for multi-objective optimal power flow considering the cost, loss, emission and voltage stability index." IET generation, transmission & distribution 6, no. 6 (2012): 515-527.
- [100]Jordehi, Ahmad Rezaee, and JasronitaJasni. "Particle swarm optimisation for discrete optimisation problems: a review." Artificial Intelligence Review 43, no. 2 (2015): 243-258.
- [101]Civicioglu, Pinar, and ErkanBesdok. "A conceptual comparison of the Cuckoo-search, particle swarm optimization, differential evolution and artificial bee colony algorithms." Artificial Intelligence Review 39, no. 4 (2013): 315-346.
- [102]Ishaque, Kashif, and Zainal Salam. "A deterministic particle swarm optimization maximum power point tracker for photovoltaic system under partial shading condition." IEEE transactions on industrial electronics 60, no. 8 (2013): 3195-3206.
- [103]Rini, Dian Palupi, SitiMariyamShamsuddin, and SitiSophiyatiYuhaniz. "Particle swarm optimization: technique, system and challenges." International Journal of Computer Applications 14, no. 1 (2011): 19-26.
- [104]Poli, Riccardo. "Analysis of the publications on the applications of particle swarm optimisation." Journal of Artificial Evolution and Applications 2008 (2008): 3.
- [105]Shi, Yuhui, and Russell Eberhart. "A modified particle swarm optimizer." In Evolutionary Computation Proceedings, 1998. IEEE World Congress on Computational Intelligence., The 1998 IEEE International Conference on, pp. 69-73. IEEE, 1998.
- [106]Reyes-Sierra, Margarita, and CA CoelloCoello. "Multi-objective particle swarm optimizers: A survey of the state-of-the-art." International journal of computational intelligence research 2, no. 3 (2006): 287-308.
- [107]Mendes, Rui, and James Kennedy. "Avoiding the pitfalls of local optima: how topologies can save the day." (2003).
- [108]Engelbrecht, Andries P. Computational intelligence: an introduction. John Wiley & Sons, 2007.
- [109]Janson, Stefan, and Martin Middendorf. "A hierarchical particle swarm optimizer." In Evolutionary Computation, 2003. CEC'03. The 2003 Congress on, vol. 2, pp. 770-776. IEEE, 2003.
- [110]Hu, Xiaohui, and Russell C. Eberhart. "Multiobjective optimization using dynamic neighborhood particle swarm optimization." (2002).
- [111]Ehrgott, Matthias. Multicriteria optimization. Springer Science & Business Media, 2006.
- [112]Reyes-Sierra, Margarita, and CA CoelloCoello. "Multi-objective particle swarm optimizers: A survey of the state-of-the-art." International journal of computational intelligence research 2, no. 3 (2006): 287-308.

- [113] Ho, S. L., Shiyong Yang, Guangzheng Ni, Edward WC Lo, and Ho-ching Chris Wong. "A particle swarm optimization-based method for multiobjective design optimizations." *IEEE Transactions on Magnetics* 41, no. 5 (2005): 1756-1759.
- [114] Tripathi, Praveen Kumar, SanghamitraBandyopadhyay, and Sankar Kumar Pal. "Multi-objective particle swarm optimization with time variant inertia and acceleration coefficients." *Information Sciences* 177, no. 22 (2007): 5033-5049.
- [115] Zhang, Guohui, Xinyu Shao, Peigen Li, and Liang Gao. "An effective hybrid particle swarm optimization algorithm for multi-objective flexible job-shop scheduling problem." *Computers & Industrial Engineering* 56, no. 4 (2009): 1309-1318.
- [116] Goh, Chi Keong, Kay Chen Tan, D. S. Liu, and Swee Chiang Chiam. "A competitive and cooperative co-evolutionary approach to multi-objective particle swarm optimization algorithm design." *European Journal of Operational Research* 202, no. 1 (2010): 42-54.
- [117] Zhang, Yong, Dun-Wei Gong, and Zhonghai Ding. "A bare-bones multi-objective particle swarm optimization algorithm for environmental/economic dispatch." *Information sciences* 192 (2012): 213-227.
- [118] Parsopoulos, Konstantinos E., and Michael N. Vrahatis. "Particle swarm optimization method for constrained optimization problems." *Intelligent Technologies—Theory and Application: New Trends in Intelligent Technologies* 76, no. 1 (2002): 214-220.
- [119] Hu, Xiaohui, and Russell Eberhart. "Solving constrained nonlinear optimization problems with particle swarm optimization." In *Proceedings of the sixth world multiconference on systemics, cybernetics and informatics*, vol. 5, pp. 203-206. 2002.
- [120] Ji, Junzhong, Hongkai Wei, and Chunnian Liu. "An artificial bee colony algorithm for learning Bayesian networks." *Soft Computing* 17, no. 6 (2013): 983-994.
- [121] Cowie, Julie, Lloyd Oteniya, and Richard Coles. "Particle swarm optimisation for learning bayesian networks." In *ICCIIS 2007, World Congress on Engineering, WCE 2007*, pp. 71-76. Newswood Limited/International Association of Engineers (IAENG), 2007.
- [122] Sahin, Ferat, M. ÇetinYavuz, ZiyaArnavut, and ÖnderUluyol. "Fault diagnosis for airplane engines using Bayesian networks and distributed particle swarm optimization." *Parallel Computing* 33, no. 2 (2007): 124-143.
- [123] Wang, Tong, and Jie Yang. "A heuristic method for learning Bayesian networks using discrete particle swarm optimization." *Knowledge and information systems* 24, no. 2 (2010): 269-281.
- [124] Valian, Maryam, and AlirezaKhanateymoori. "A Method for structural Learning in Bayesian Networks based on Particle Swarm Optimization."
- [125] Chun-Feng, Wang, and Liu Kui. "A Novel Hybrid Method for Learning Bayesian Network."
- [126] Fortier, Nathan, John Sheppard, and Shane Strasser. "Learning bayesian classifiers using overlapping swarm intelligence." In *Swarm Intelligence (SIS), 2014 IEEE Symposium on*, pp. 1-8. IEEE, 2014.
- [127] Microsoft Visual Studio Professional 2015, Version 14.0.25123.00 Update 2, Microsoft .NET Framework, Version 4.6.01055, Installed Version: Professional
- [128] Microsoft SQL Server Management Studio, 12.0.2000.8, SQL Server 2014
- [129] Marco Scutari, Ph.D. ,BazesianNetwor repository, <http://www.bnlearn.com/>
- [130] Gentleman, Robert, Ross Ihaka, and D. Bates. "The R project for statistical computing." URL: <http://www.r-project.org/254> (2009).
- [131] Ripley, Brian D. "The R project in statistical computing." *MSOR Connections. The newsletter of the LTSN Maths, Stats & OR Network* 1, no. 1 (2001): 23-25.