37. Саветовање CIGRE Србија 2025
СИГУРНОСТ, СТАБИЛНОСТ, ПОУЗДАНОСТ И RESILIENCE ЕЛЕКТРОЕНЕРГЕТСКОГ СИСТЕМА
МУЛТИСЕКТОРСКО ПОВЕЗИВАЊЕ У ЕНЕРГЕТИЦИ И ПРИВРЕДИ
26 - 30. мај 2025, Копаоник

DOI: 10.46793/CIGRE37.C2.21

C2.21

PRIMENA EXTENDED ISOLATION FOREST ALGORITMA ZA DETEKCIJU ANOMALIJA U PODACIMA ELEKTROENEREGETSKOG SISTEMA

APPLICATION OF THE EXTENDED ISOLATION FOREST ALGORITHM FOR ANOMALY DETECTION IN POWER SYSTEM DATA

Vladimir Bečejac[12] Dobrila Škatarić[1], Petar Lukić[1], Dragan Ćetenović[3]

Kratak sadržaj: Detekcija anomalija u elektroenergetskim sistemima važna je za pouzdanost i stabilnost mreže, a posebno u kontekstu integracije obnovljivih izvora energije i sve veće složenosti strukture elektroenergetskog sistema. U ovom radu se istražuje primena Extended Isolation Forest (EIF) algoritma za detekciju anomalija u raznim merenjima koje operatori prenosnih sistema dobijaju u svakodnevnom radu. EIF predstavlja poboljšanu verziju standardnog Isolation Forest algoritma, omogućavajući bolju obradu podataka visoke dimenzionalnosti i veće preciznosti u otkrivanju odstupanja. Rezultati istraživanja ukazuju na sposobnost da EIF precizno otkrije anomalije u realnim i simuliranim podacima. Prikazani pristup može doprineti razvoju naprednih sistema za online nadzor i prevenciju poremećaja u elektroenergetskim sistemima.

Ključne reči: detekcija anomaija, pouzdanost, stabilnost, extended isolation forest

1 UVOD

Detekcija anomalija je oblast mašinskog učenja koja se bavi identifikacijom podataka koji odstupaju od većine ostalih podataka. Pronalaženjem ovih anomalija moguće je unaprediti kvalitet podataka njihovim uklanjanjem, prepoznati kada sistem ulazi u retka i/ili potencijalno opasna stanja, kao i pravovremeno detektovati neuobičajene obrasce koji mogu ukazivati na sajber-napade ili kvarove unutar mreže elektroenergetskog sistema.

Razvoj tehnologije mernih uređaja, posebno sinhrofazorskih uređaja (PMU – Phasor Measurement Units), omogućio je prikupljanje visoko preciznih, sinhronizovanih i detaljnih informacija o stanju elektroenergetskog sistema [1]. Sve veća primena PMU uređaja doprinela je povećanju pouzdanosti i efikasnosti upravljanja sistemima [2] kroz praćenje ključnih parametara kao što su stabilnost napona [3], ugaona stabilnost [4], oscilacije [5], detekcija mesta kvarova u elektroenergetskom sistemu [6]. U Srbiji su prvih 6 PMU uređaja je instalirano 2015. godine [7], a danas ih je 32, sa planom da se u 2025. godini instalira još 7. Ovo je dodatno intenziviralo istraživanja u oblasti njihove primene.

[1] Mašinski fakultet – Univerzitet u Beogradu
[2] AD Elektromreža Srbije
[3] Fakultet tehničkih nauka Čačak, Univerzitet u Kragujevcu

Kako količina dostupnih PMU podataka brzo raste, primena metoda za analizu podataka i rudarenje podataka (Data Mining) postala je ključna za operatere u praćenju dinamike elektroenergetskog sistema i pravovremenom otkrivanju stanja koja bi mogla dovesti do havarija ili raspada sistema. Svedoci smo da su se u samo poslednje 4 godine dogodila dva ozbiljna incidenta u evropskoj interkonekciji i to razdvajanje interkonekcije 8. januara 2021. godine i blackout 21. juna 2024. godine. U obe situacije, na osnovu PMU podataka operatori prenosnih sistema su mogli da nemile događaje preduprede. Različite tehnike, uključujući analizu sopstvenih vrednosti (Eigenvalue Analysis - EA), Furijeovu transformaciju (Fourier Transform - FT), kratkotrajnu Furijeovu transformaciju (Short Time Fourier Transform - STFT) i spektralnu analizu (Spectral Analysis - SA), istraživane su u svrhu detekcije anomalija [8, 9,10]. Međutim, zbog nestacionarnosti PMU signala, tradicionalni pristupi često nisu dovoljno efikasni ili zahtevaju velike računarske resurse, što dovodi do potrebe za razvojem novih metoda koje su prilagođene specifičnostima PMU podataka.

Jedna od poznatih metoda detekcije anomalija koja može prevazići ove izazove je Isolation Forest (IF), algoritam zasnovan na konceptu izolacije, odnosno merenju koliko je lako izdvojiti određeni podatak od ostalih. Tokom godina, razvijeni su algoritmi koji dodatno unapređuju osnovni IF, poput Extended Isolation Forest (EIF) i Hybrid Isolation Forest (HIF) [11,12]. EIF menja način na koji se vrše grananja unutar algoritma, pružajući preciznije ocene anomalija, dok HIF uvodi izračunavanje centroida, omogućavajući bolju detekciju složenijih oblika anomalija.

U ovom radu ispituje se primena algoritma Extended Isolation Forest za detekciju anomalija specifično u podacima elektroenergetskog sistema. Cilj rada je da prikaže mogućnosti i prednosti EIF algoritma u realnom elektroenergetskom okruženju.

Rad je organizovan na sledeći način. U drugoj glavi su opisane anomalije u opštem smislu i najčešće korišćene metode za njihovu detekciju. U trećoj glavi je objašnjen Isolation Forest algoritam sa svojim prednostima i manama. Zbog ispoljenih mana, uveden je Extended Isolation Forest i objašnjen je u glavi četiri. Peta glava je rezervisana za rezultate i komentare proračuna i to za dva ozbiljna događaja u interkonekciji Kontinentalne Evrope. U šestoj glavi je napisan zaključak rada.

2 ANOMALIJE I METODE DETEKCIJE

Identifikacija anomalija je od ključnog značaja u različitim oblastima zbog potencijala da ukaže na retke, ali važne pojave ili stanja. Motivacija za detekciju anomalija najčešće proizilazi iz potrebe da se poboljša tačnost modela, uklone nevažeći, netačni ili maliciozni podaci, identifikuju nove vredne informacije ili prilagode postojeći modeli kako bi bolje odgovarali realnim uslovima.

Anomalije mogu ukazivati na kritične situacije kao što su kvarovi sistema, neželjeni uticaji ili namerne zloupotrebe, pa njihovo pravovremeno otkrivanje omogućava pravovremene intervencije.

Važno je razlikovati pojmove anomalija jer njihova uloga u analizi podataka nije ista:

- Anomalija (engl. anomaly ili outlier) predstavlja tačku koja značajno odstupa od većine ostalih podataka i obično ukazuje na posebna ili neočekivana stanja.

- Šum (engl. noise) odnosi se na neželjene ili slučajne podatke koji nemaju vrednost za analizu i najčešće se uklanjaju pre daljih analiza.

- Novost (engl. novelty) prvobitno se javlja kao anomalija, ali nakon stručne procene postaje deo regularnih podataka i koristi se za ažuriranje postojećeg modela.

Anomalije se mogu klasifikovati na sledeći način:

- Tačkaste anomalije: izolovane tačke koje jasno odstupaju od većine podataka.

- Kontekstualne anomalije: odstupanja koja su anomalna samo u specifičnom kontekstu, na primer neuobičajene vrednosti u određenim vremenskim periodima.

- Kolektivne anomalije: grupe podataka koje pojedinačno nisu anomalne, ali predstavljaju anomaliju kada se analiziraju zajedno.

Ovaj rad fokusira se isključivo na tačkaste anomalije.

Detekcija anomalija suočava se sa određenim specifičnim problemima, kao što su:

- Maskiranje: situacija u kojoj blizina drugih anomalija sprečava identifikaciju neke anomalije.

- Utapanje: pogrešno prepoznavanje normalnih podataka kao anomalija usled njihove blizine pravim anomalijama.

Algoritmi se mogu podeliti u tri kategorije prema tipu podataka:

- Supervizovano učenje: dostupne oznake i za normalne podatke i za anomalije (nije predmet ovog rada).

- Polusupervizovano učenje: dostupne oznake samo za jednu kategoriju, obično normalne podatke.

- Nesupervizovano učenje: bez dostupnih oznaka; najčešće korišćena tehnika zbog pretpostavke da su anomalije retke u odnosu na normalne podatke.

Extended Isolation Forest, koji je predmet analize ovog rada, spada u kategoriju nesupervizovanih metoda.

Izlaz algoritama može biti binaran (normalan ili anomalija) ili kontinuiran u obliku ocena (score) anomalija koji se kreće između 0 i 1. Korisnik postavlja prag na osnovu kog odlučuje koje tačke su anomalije.

Algoritmi za detekciju anomalija mogu biti zasnovani na različitim pristupima, poput klasifikacije, klasterovanja, statističkih metoda ili neuronskih mreža. Extended Isolation Forest, kao metoda zasnovana na šumama odluka (decision tree), predstavlja zasebnu kategoriju – „ensemble-based“ metode, koje koriste princip izolacije tačaka za efikasnu identifikaciju anomalija.

Ne postoji jedinstvena, opšteprihvaćena, formalna definicija anomalije. Stoga se u literaturi može pronaći kolekcija više različitih definicija anomalija.

Jedna je definicija anomalije od strane Hawkinsa koji kaže: „Anomalija je posmatranje koje *toliko odstupa od drugih posmatranja da izaziva sumnju da je generisano drugačijim mehanizmom."* Primetiti da ovo nije rigorozna niti matematička definicija. Do sada su date mnoge definicije i metode koje su matematički utemeljenije od pomenute. Biće pomenute metode koje detektuju anomalije na osnovu udaljenosti, neparametarske gustine, parametarske gustine, klastera, dubine i izolacije.

U oblasti detekcije anomalija, Mahalanobisova udaljenost se koristi za identifikaciju višedimenzionalnih autsajdera analizirajući udaljenost između svake tačke i srednje vrednosti uzorka uz korekciju prema matrici kovarijance. Neparametarske metode zasnovane su na gustini i fokus je na lokalnim regionima skupa/skupova podataka, gde anomalije predstavljaju tačke sa manjom lokalnom gustinom kada se uporede sa njihovim okruženjem. Sve ove metode ne zavise direktno od predpostavke o distribuciji podataka, čime omogućavaju veliku fleksibilnost u praktičnoj primeni.

Parametarske metode, s druge strane, zasnivaju se na specifičnoj distribuciji gustine podataka i koriste parametre te distribucije za identifikaciju anomalija. Metode zasnovane na klasterima, kao što je *K*-means, vrše podele podataka na klastere i identifikuju anomalije kao one tačke koje ne pripadaju nijednom klasteru. Za ocenu efikasnosti ovih metoda koristi se matrica klasifikacije, koja pruža uvid u performanse algoritma na osnovu četiri moguća ishoda klasifikacije.

3 ISOLATION FOREST ALGORITAM

Kako je već naglašeno, anomalije su tačke koje odstupaju od drugih tačaka. Kako ima manje anomalija nego regularnih tačaka, podložnije su izolaciji. Algoritmi Isolation Forest i njegova nadogranja, koji su glavna tema ovog rada, zasnivju se na principu izolacije. Metod Isolation Forest ne stvara profil normalnih instanci, već direktno izoluje anomalije. U osnovi IF metode je stablo izolacije, koje je pravilno binarno stablo. To znači da svaki čvor u stablu ima tačno nula ili dva čvora potomka. Ovde se pretpostavlja da je T čvor stabla izolacije. Spoljni čvorovi nemaju čvorove potomke, a unutrašnji čvorovi imaju tačno dva čvora potomka (TL, TR).

Pretpostavimo da postoji skup podataka *X* od *n* instanci iz *d*-varijantne distribucije i da se na njemu koristi IF. Tada se *X* rekurzivno deli slučajnim izborom karakteristike $Q_i \in Q_1, Q_2, \dots, Q_d$ sa jednakom verovatnoćom iz skupa karakteristika i vrednošću deljenja p koja daje stablo izolacije. Nakon što je izabrana karakteristika Q_i, pripadajuća vrednost ovoj specifičnoj karakteristici $X(Q_i)$ se upoređuje sa vrednošću deljenja p za svaku tačku podataka. Ako je $X(Q_i) < p$, tačka podataka će ići u TL, u suprotnom će ići u TR. Ovo se nastavlja dok stablo ne dostigne limit visine (jer se drvo seče na unapred podešenoj granici visine kako bi se smanjilo vreme izvršavanja IF algoritma. U suprotnom detekcija anomalija bi u određenim situacijama isuviše dragocenog vremena oduzela.) ili dok u skupu *X* ne ostane samo jedna tačka podataka ili sve podatke u *X* imaju iste vrednosti. Ovo je prva faza modela i naziva se faza obuke. U ovoj fazi, stabla izolacije se konstruišu iz poduzorka podataka.

U stablima izolacije, instance se rekurzivno dele dok sve ne budu izolovane. Anomalije se izoluju ranije u stablima nego normalne instance, zbog njihovih razlikovnih vrednosti atributa.

Druga faza je faza evaluacije, gde se izvodi ocena anomalije s iz očekivane dužine puta $E\big(h(x)\big)$ za svaku instancu. Dužina puta $h(x)$ tačke podataka x predstavljena je brojem ivica od korena do završnog čvora dok tačka x prolazi kroz stablo izolacije. Očekivana dužina puta izvodi se prolaskom svih tačaka podataka kroz svako stablo izolacije u šumi izolacije.

Tada je to prosečna vrednost $h(x)$ iz svih stabala izolacije koja su izgrađena. Međutim, stabla imaju limit visine i može se desiti da stablo nije potpuno izraslo. To su čvorovi sa ranim prekidom, što znači da će ti čvorovi sadržati više od jedne tačke podataka. Ako je to slučaj, dodatna konstanta $c(n)$ se dodaje dužini puta instance u čvoru sa ranim prekidom. Ova konstanta $c(n)$ je prosečna dužina puta stabla izolacije koje je izgrađeno sa n tačaka podataka. Konačno, ocena anomalije (skor) tačke podataka *x* za skup podataka veličine *n* daje se kao:

$$s(x,n) = 2^{-\frac{E(h(x))}{c(n)}}, \tag{1}$$

gde je $E(h(x))$ prosečna vrednost $h(x)$ iz svih stabala izolacije koja su izgrađena. Prosečna vrednost $h(x)$ neuspešne pretrage u binarnom stablu za skup podataka veličine *i* je data formulom:

$$c(i) = \begin{cases} 2H(i-1) - \frac{2\cdot(i-1)}{i}, & i > 2 \\ 1, & i = 2 \\ 0 & \text{inače} \end{cases}, \tag{2}$$

gde je *H* harmonijski broj, procenjen kao $ln(.) + 0.5772$ (Ojlerova (Euler) konstanta).

Razlikujemo tri slučaja:

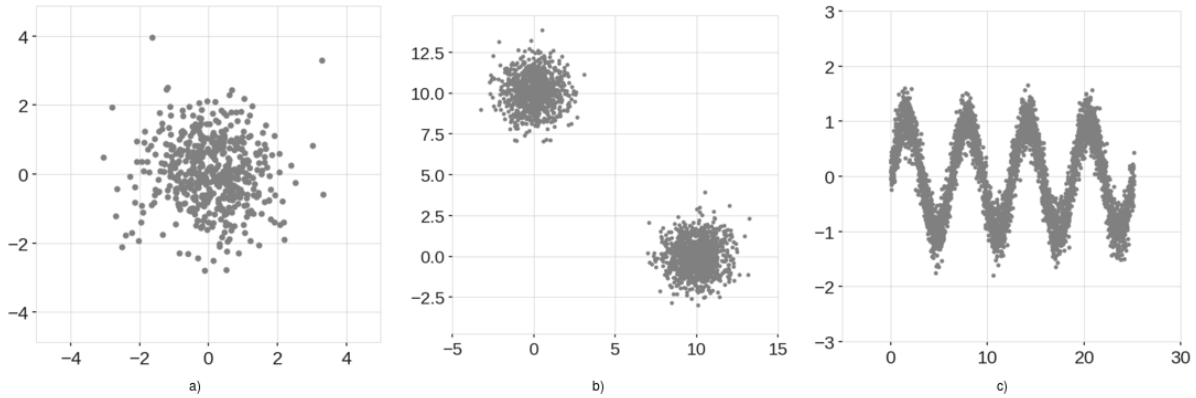
- kada je $E(h(x)) \longrightarrow c(n)$: $s(x,n) = 2^{-\frac{c(n)}{c(n)}} = 2^{-1} = 0.5$;
- kada je $E(h(x)) = 0$: $s(x,n) = 2^0 = 1$;
- kada je $E(h(x)) \longrightarrow n-1$:

$$\begin{aligned} s(x,n) &= 2^{-\frac{n-1}{c(n)}} \\ &= 2^{-\frac{1}{2}^{\frac{n-1}{\ln n-1+0.57722-1}}} \\ &\longrightarrow 0,\ n \longrightarrow +\infty. \end{aligned} \tag{3}$$

Ukoliko je skor $s(x)$ blizu 1, tada je podatak *x* anomalija, a ukoliko je $s(x)$ manji od 0.5, može se smatrati da je podatak regularan. Ukoliko se skor nalazi u određenoj epsilon okolini vrednosti 0.5, tada ne možemo sa sigurnošću da kažemo da li je anomalija ili regularan podatak.

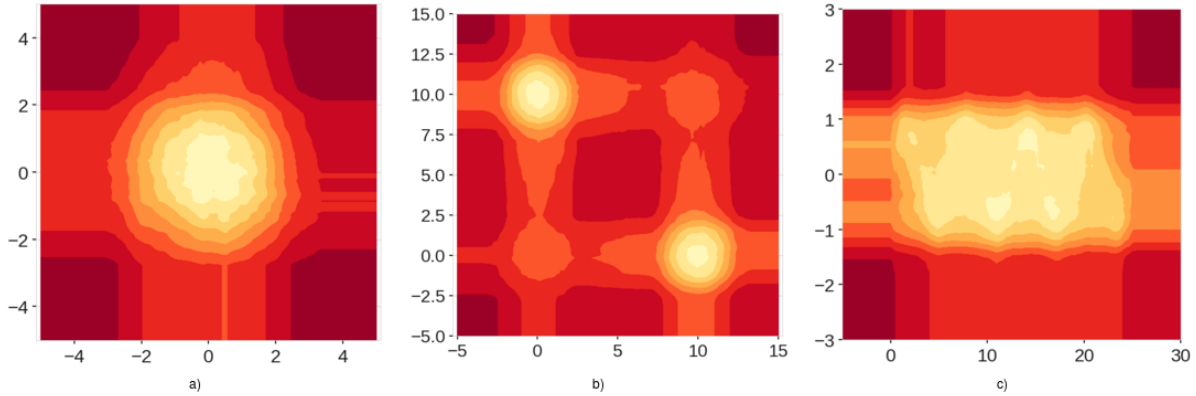
IF koristi minimalne potrebe za memorijom i računanjem. Treba spomenuiti i skor anomalije, koji pati od pristrasnosti uzrokovane načinom na koji se formira stablo. U radu [13] se predlaže jednostavno i ingeniozno rešenje za otkalanjanje ove pristrasnosti i izgradnju robustnijeg modela sa istim performansama. IF metoda može lako raditi sa podacima velikih dimenzija, ne zavisi od modela i brzo se skalira.

Međutim, gledajući grafike skorova anomalija za neke jednostavne primere, možemo videti da skorovi anomalija proizvedeni standardnom šumom izolacije nisu konzistentni. Ukoliko se posmatraju tri primera sa slike 1 možemo primetiti da se regularne tačke nalaze oko koordinatnog početka na a) slici. Na slici b) mogu se primetiti dve regularne grupe podataka, dok na slici c) regularni podaci se pokoravaju sinusnom zakonu.



Slika 1: Posmatrani primeri

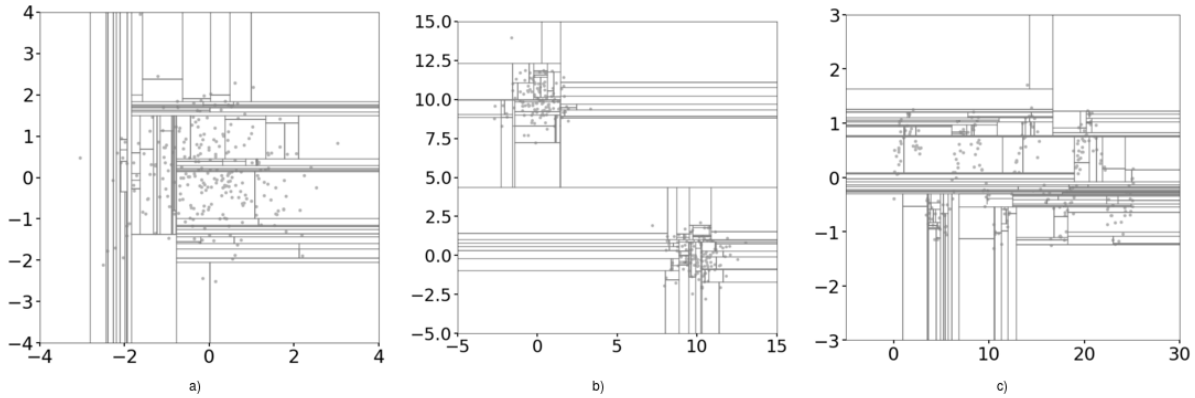
U svakom slučaju, koristimo podatke za obuku naše šume izolacije. Zatim koristimo obučene modele da ocenimo kvadratnu mrežu uniformno raspoređenih tačaka podataka, što rezultira mapama skorova prikazanim na slici 2. Na primer, za podatke prikazane na slici 1a, očekujemo da vidimo niske skorove anomalija u centru mape, dok bi skor anomalija trebao rasti kako se udaljavamo radijalno od centra. Slično za ostale figure. Gledajući na mape skorova proizvedene standardnom šumom izolacije prikazane na slici 2, jasno možemo videti nekonzistentnosti u skorovima. Dok jasno možemo videti region niskog skora anomalija u centru na slici 2a, takođe možemo videti regione poravnate sa x i y osama koje prolaze kroz početak koji imaju niže skorove anomalija u poređenju sa četiri ugla regiona. Na osnovu našeg intuitivnog razumevanja podataka, ovo ne može biti tačno. Slična pojava se primećuje na slici 2b. U ovom slučaju, problem je pojačan. Pošto postoje dva klastera, veštački niski regioni skora anomalija se ukrštaju blizu tačaka (0,0) i (10,10) i stvaraju niske regione skora anomalija gde nema podataka. Odmah je očigledno kako ovo može biti problematično. Što se tiče trećeg primera, slika 2c pokazuje da je struktura podataka potpuno izgubljena. Sinusoidalni oblik se u suštini tretira kao jedan pravougaonik, što je naravno pogrešno.



Slika 2: Ocena anomalija (skor) za IF

Ispostavlja se da je proces razdvajanja, opisan iznad, glavni izvor pristrasnosti uočen na mapama skorova. Slika 3 pokazuje proces opisan iznad za svaki od razmatranih primera. Rezovi grana su uvek paralelni sa osama, i kao rezultat prekomerne izgradnje mnogih drveta, regioni u domenu koji ne zauzimaju tačke podataka dobijaju suvišne rezove grana.

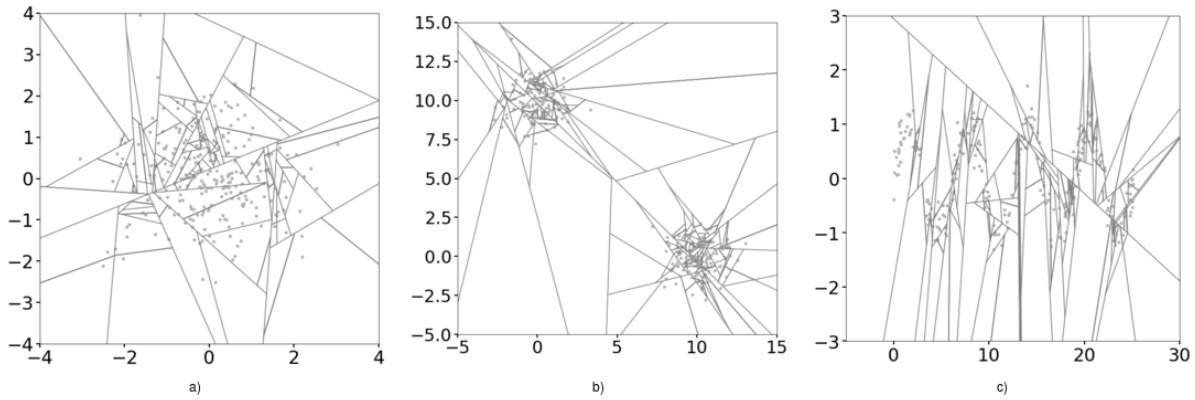
Zbog ovoga se uvodi Extended Isolation Forest (EIF) metoda o kojoj će biti reči u sledećem poglavlju.



Slika 3: Proces biranja rezova za IF

4 EXTENDED ISOLATION FOREST

Proširena šuma izolacije rešava pomenuti problem omogućavajući da se proces grananja odvija u svim pravcima. Proces biranja rezova grana se menja tako da se na svakom čvoru, umesto biranja slučajne karakteristike zajedno sa slučajnom vrednošću, biramo slučajni normalni vektor zajedno sa slučajnom tačkom preseka. Prosto, rezovi nisu paralelni sa x i y -osom, već su prikazani na slici 4.



Slika 4: Proces biranja rezova za EIF

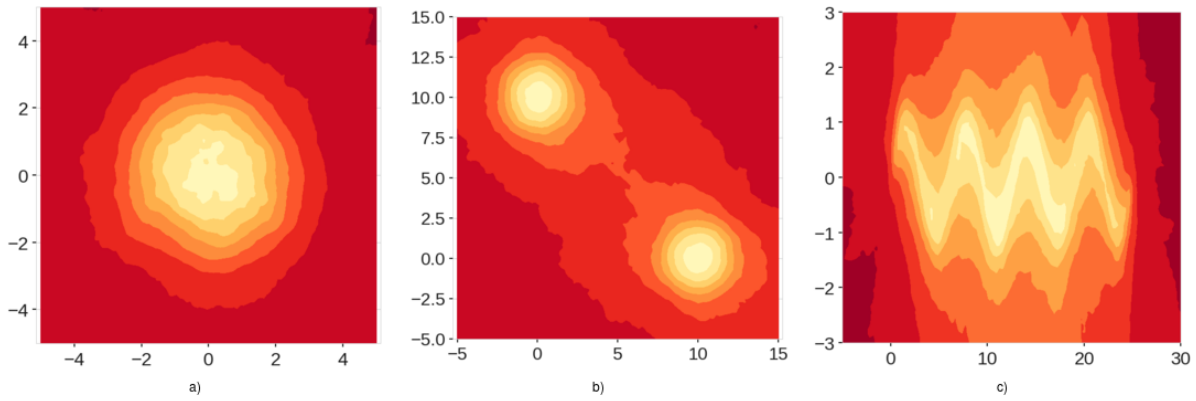
Ova verzija algoritma uvodi korišćenje hiper-ravni sa nasumičnim nagibima za deljenje podataka, umesto standardnog vertikalnog ili horizontalnog rezanja. Ovo smanjuje pristrasnost i artefakte koji mogu nastati tokom procesa dodeljivanja anomalijskih skorova i omogućava pouzdanije mapiranje skorova.

EIF koristi nasumičnost u izboru karakteristika i vrednosti za efikasnu izolaciju anomalijskih tačaka, koje se brzo ističu zahvaljujući ovim nasumičnim izborima. Algoritam zahteva dve ključne informacije za svaki proces grananja: nasumični nagib preseka i nasumični presek, koji se bira iz dostupnog opsega vrednosti trening podataka.

Za N-dimenzionalni skup podataka, izbor nasumičnog nagiba preseka podrazumeva odabir normalnog vektora, jednako distribuiranog duž N-sfere, što se postiže generisanjem nasumičnog broja za svaku koordinatu vektora iz normalne raspodele. Ovo omogućava da presečne tačke budu izabrane nasumično i pravično, bez dodatne pristrasnosti koja može nastati zbog fiksne orijentacije preseka.

Ovaj pristup omogućava znatno poboljšanje u otkrivanju anomalija, posebno na ivicama skupova podataka, gde je važno precizno identifikovanje izuzetaka. EIF efikasno rešava probleme koje standardni IF ne može da adresira, omogućavajući kreiranje uniformnih i bezpristrasnih mapa skorova anomalija. Ova poboljšanja postižu se bez žrtvovanja efikasnosti izračunavanja, čineći EIF izuzetno pogodnim za praktičnu primenu u analizi anomalija.

Nakon korišćenja EFI algortima dobijaju se skorovi anomalija kao na slici 5. Jasno se može uočiti da uvedene modifikacije rešavaju problem sa mapama skorova i prozivode pouzdane rezultate.



Slika 5: Ocena anomalija (skor) za EIF

5 REZULTATI

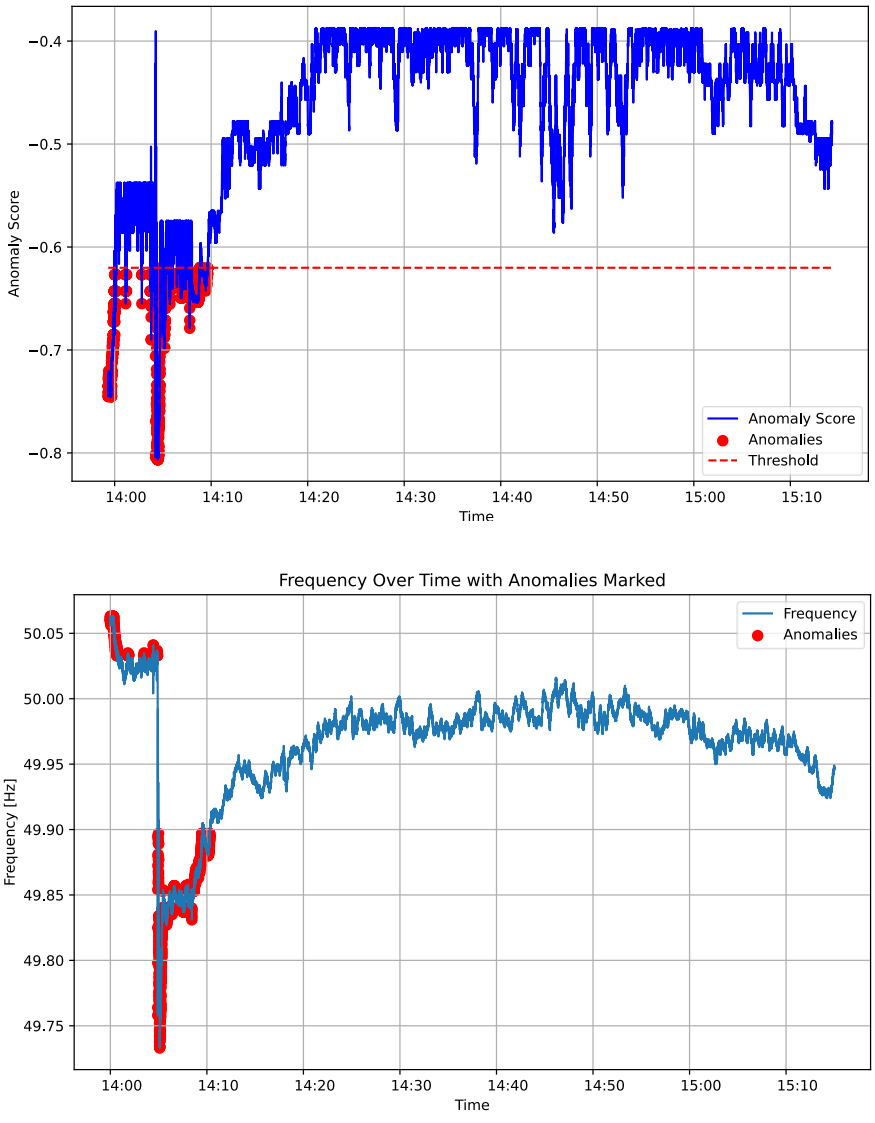
Za potrebe primene EIF algoritma za detekciju anomalija, razvijen je programski kod u jeziku Python. Koristite standardne biblioteke kao što su `pandas` za rad sa podacima, matplotlib.pyplot za vizualizaciju, numpy za numeričke operacije, i sklearn.ensemble za Isolation Forest model, koji smo modifikovali.

Korisnik otprema Excel fajl koji sadrži podatke. Pretpostavlja se da fajl sadrži kolonu Frekvencija [Hz] i kolonu Vreme [s] za indeksiranje. Broj nasumičnih nagiba u rezovima je postavljen na 2. Testirano je sa više stabala u šumi i dobijeno je da za ovu vrstu analize je sasvim zadovoljavajući broj 100. Veći broj stabala poboljšava preciznost detekcije, ali na drugu stranu oduzima više vremena. Maksimalni broj uzoraka koji se koristi za gradnju svakog stabla je $n = min\{256, n_r\}$, gde je sa n_r označen broj redova u podacima. Kontaminacija je postavljena na 5%, što znači da očekujemo da će se u podacima pojaviti ovoliko anomalija od ukupnog broja.

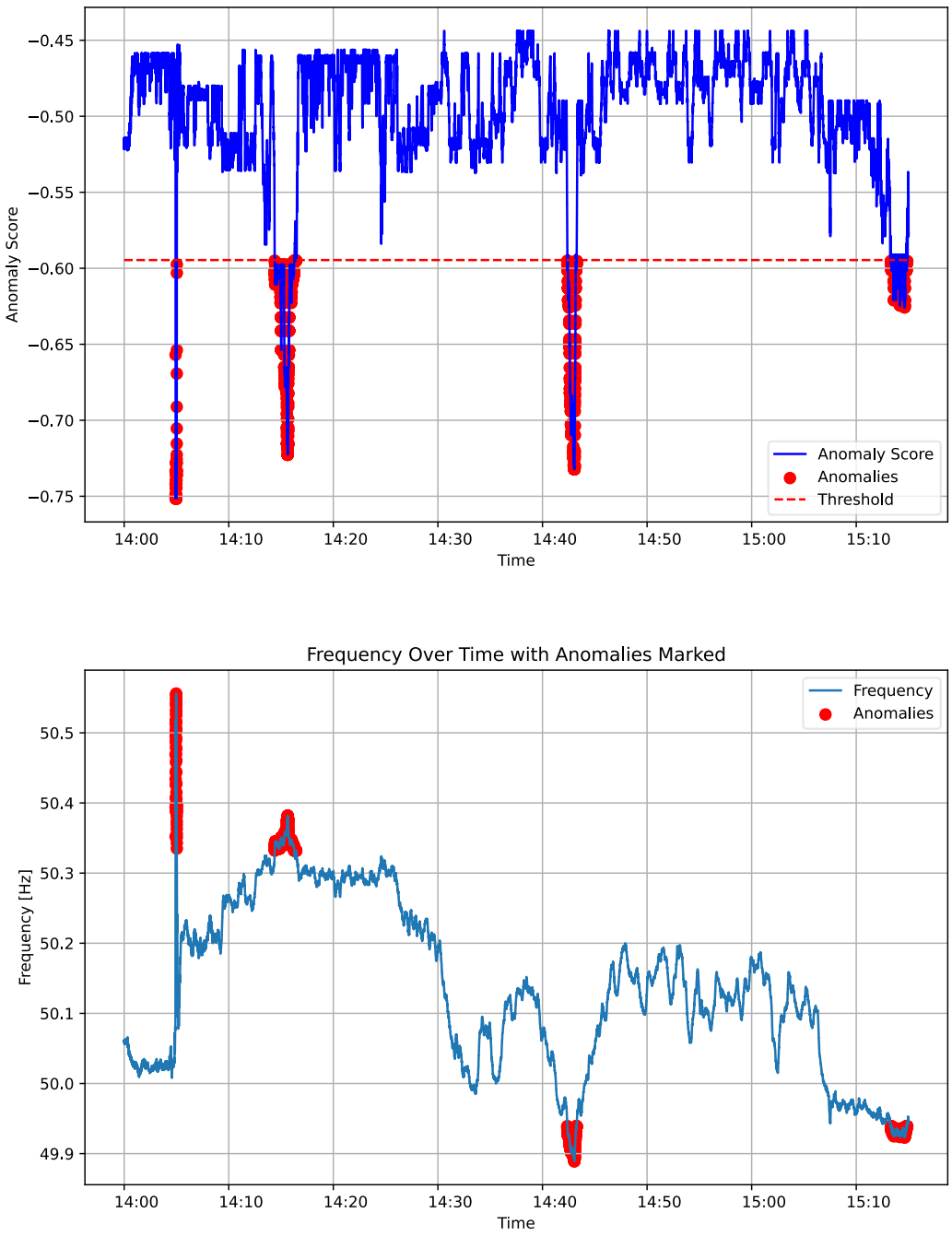
U ovom odeljku biće prikazani rezultati navedenih proračuna. Uzeto je nekoliko realnih primera koji su se dogodili u poslednjih par godina u interkonekciji Kontinentalne Evrope [14,15]. Dva događaja su javnosti već opšte poznata. Jedan je razdvajanje interkonekcija na dva dela 8. januara 2021. godine, a drugi je blackout zemalja na Balkanu (deo Crne Gore i Hrvatske i potpuno Bosna i Hercegovina i Albanija).

Treći događaj je manji sistemski poremećaj, koji se dogodio 9. maja 2024. godine kada je u 09:17 usled ispada 400 kV dalekovoda Pakš – Šandorfalva (opterećen u tom trenutku 1500 MW) došlo do naglog povećanja opterećenja po dalekovodima drugih sistema i to 400 kV Žerjavinec – Ernestinovo i 400 kV Maribor – Cirkovice.

Prvi analizirani dođaj desio se 8. januara 2021. Na slici 6, je prikazana frekvencija na vremenskom intervalu od 14:00:00 do 15:15:00 za trafostanicu Ernestinovo. Gornja slika ilustruje ocene anomalija i svaka je označena crvenom bojom. Regularne tačke su plave boje. Na donjoj slici je prikazana frekvencija u funkciji vremena, sa crvenim anomalnim tačkama. Usled lošeg uklopnog stanja u ovoj trafostanici, ispadom prekidača u spojnom polju 400 kV od prekostrujne zaštite, dolazi do kaskadnih događaja koji su za reperkusiju imali razdvajanje interkonekcije. Sa slike 6 možemo zaključiti da se značajne vrednosti anomalija pojavljuju u 14:04, što koncidira sa vremenom kada je došlo do razdvajanja. Međutim, ovo ne mora označavati ozbiljan sistemski poremećaj. Zbog toga se paraleleno moraju analizirati podaci sa više sinhrofazorskih uređaja. Na slici 7 je prikazan grafik frekvencije u istom vremenskom intervalu, ali sada iz TS Solun u Grčkoj. Tu se jasno mogu videti anomalije u istom vremenskom trenutku. Kako je frekvencija sistemski parametar, kako se anomalije dešavaju u istim trenucima i kako je frekvencija u TS Ernestinovo naglo skočila, a u TS Solun naglo pala, ove anomalije sugerišu operatorima da nisu u pitanju greške u merenjima već ozbiljan sistemski poremećaj.



Slika 6: Detektovane anomalije u TS Ernestinovo



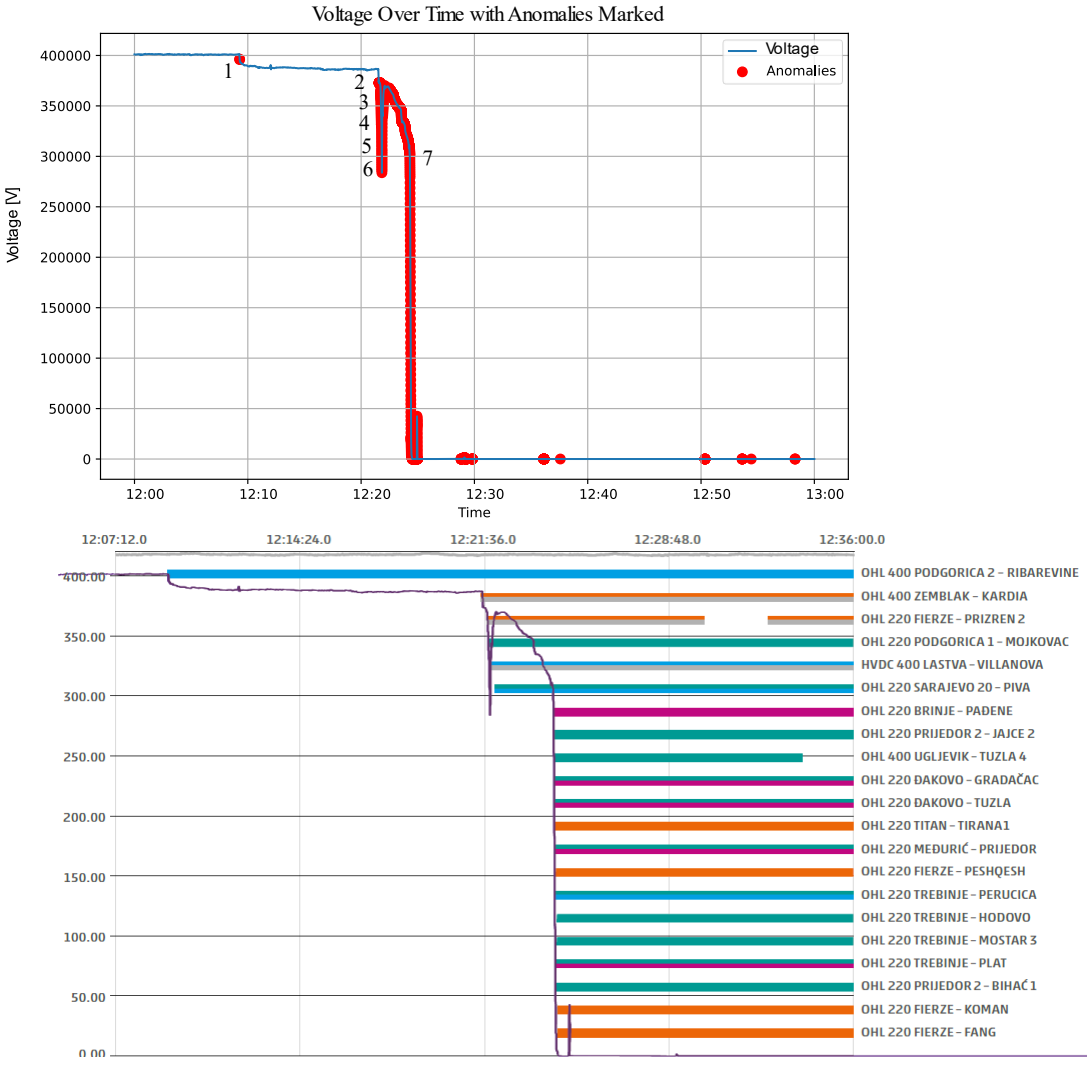
Slika 7: Detektovane anomalije u TS Solun

Sledeći analizirani događaj je raspad elektroenergetskih sistema na Balkanu (blackout) koji se dogodio 21. juna 2024. godine. Na gornjem delu slike 8 je prikazan grafik 400 kV napona u TS Trebinje (NOS BIH).

Posledica prekida 400 kV dalekovoda Podgorica 2 - Ribarevine je pad napona u Crnoj Gori, Bosni i Hercegovini i Dalmaciji. Na slici 8 je prikazano tačkom 1 i mogu se videti tada detektovane anomalije.

Ovaj ispad je dalje izazvao promenu položaja regulacionih transformatora (za 100/x kV/kV to je bila uvek promena na višu poziciju, što znači da se pri smanjenom naponu na strani visokog napona pokušava održati isti napon na strani niskog napona), i situacija se stabilizovala u novom stacionarnom stanju. U tom periodu, od tačke 1 do tačke 2 se i nisu detektovale anomalije, što je i dobro jer se desila prirodna reakcija sistema.

Tačkom 2 na slici je označena ispad 400 kV interkonektivnog dalekovoda Zemblak - Kardia, između Albanije i Grčke. Posledica ovog prekida je pad napona u istim kontrlonim oblastima koje su pomenute. Kvar na MONITA kablu (Crna Gora - Italija) kratkotrajno je povećao lokalne napone i to je označeno tačkom 5. Primetiti da razvijeni algoritam opravdano sve vreme detektuje anomalije.

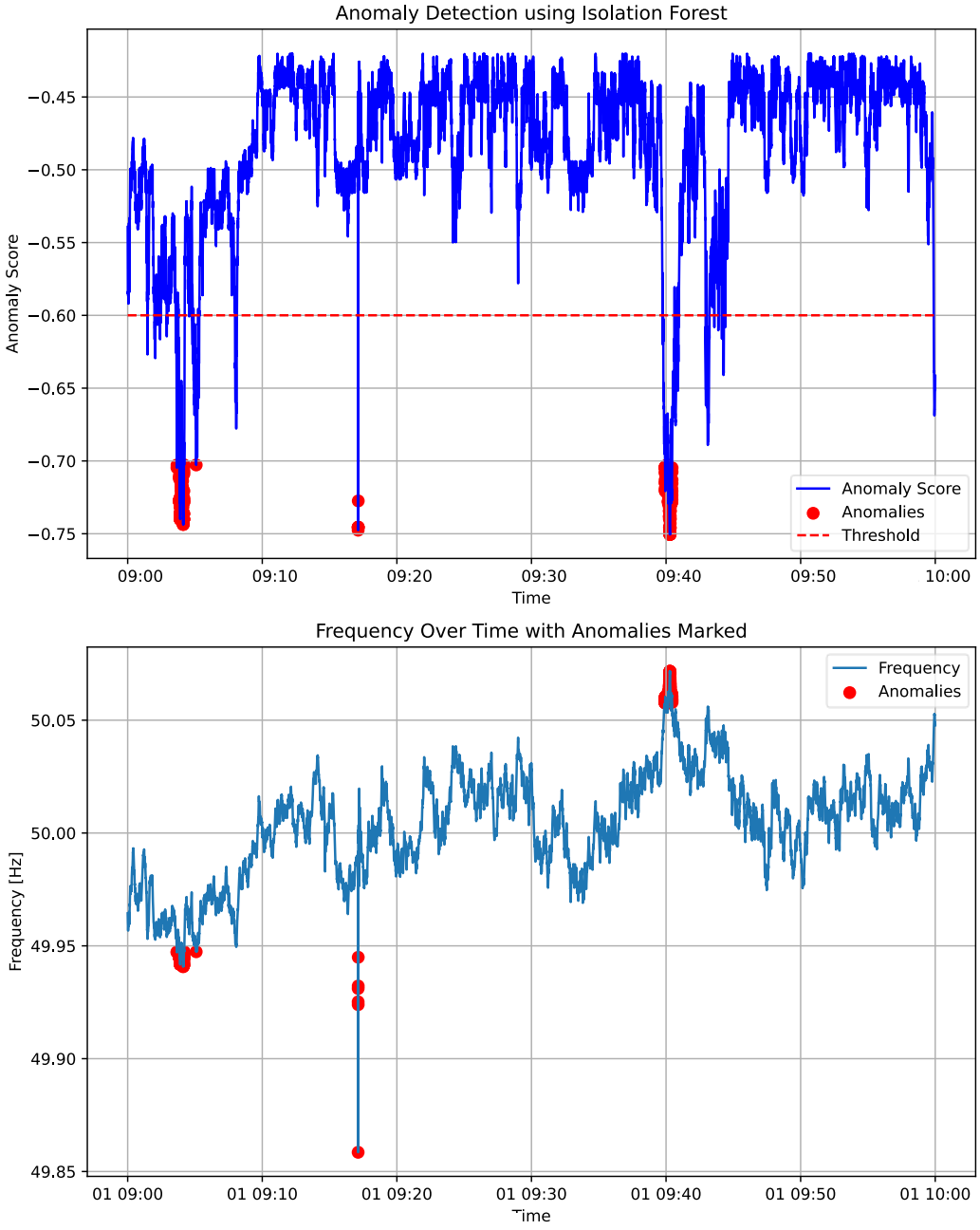


Slika 8: Detektovane anomalije u TS Trebinje

U 3 – 4 sekunde, pre trenutka bez napona u svim kontrolnim oblastima, došlo je do naglog pada napona i to je označeno tačkom 7 na gornjem grafiku na slici 8. Preklopnici transformatora u trafostanicama gde je primenjena automatska regulacija napona i koji prethodno nisu bili na krajnjem koraku automatski su reagovali, ali zbog kratkog perioda to se nije desilo u svim trafostanicama. Nakon tačke 7 detektovane anomalije nemaju smisla jer je trafostanica ostala bez napona i eventualne promene na PMU uređaju su posledica greške (u merenju, šumu itd.).

Donji grafik na slici 8 predstavlja listu ispada dalekovoda i kako su oni uticali na napon u TS Trebinje.

Na slici 9 je prikazana frekvencija za treći analizirani događaj od 9. maja 2024. godine. Prikazani podaci su sa PMU uređaja dobijenih iz TS Ernestinovo, koja je bila značajna trafostanica tokom incidenta jer je nakon ispada 400 kV dalekovoda Pakš – Šandorfalva, on preuzeo veliki deo opterećenja. Možemo primetiti da su detektovane anomalije u kritičnim trenucima



Slika 9: Detektovane anomalije u TS Ernestinovo

Iz prethodne analize možemo zaključiti da algoritam dobro detektuje anomalije kako za velike, tako i za manje sistemske poremećaje i da je kontaminacija postavljena na sasvim prihvatljiv nivo.

Ovakva primena EIF algoritma pokazuje da se anomalije mogu pouzdano detektovati i u slučajevima koji ne dovode do potpunog razdvajanja sistema, već samo do lokalizovanih promena tokova snaga. Na osnovu ovakvih rezultata može se predložiti integracija EIF algoritma u sistem ranog upozoravanja dispečera, gde bi se detekcija manjih anomalija u realnom vremenu koristila kao indikator za preventivne korektivne mere, čime bi se unapredila stabilnost i sigurnost elektroenergetskog sistema.

6 ZAKLJUČAK

Primena Extended Isolation Forest algoritma pokazala je značajne prednosti u detekciji anomalija unutar elektroenergetskih sistema, koristeći podatke iz sinhronizovanih fazorskih mernih jedinica (PMU). Kroz poboljšanja u odnosu na standardni Isolation Forest algoritam, EIF omogućava efikasniju obradu podataka visoke dimenzionalnosti, što rezultira preciznijim otkrivanjem odstupanja od normalnog ponašanja sistema. Ova sposobnost je posebno korisna u kontekstu sve veće kompleksnosti elektroenergetskih sistema i integracije obnovljivih izvora energije.

EIF algoritam, kroz svoje unapređene mehanizme grananja i prilagođavanje za visoko-dimenzionalne podatke, pokazuje bolju sposobnost u identifikaciji kako očitih tako i suptilnijih anomalija koje mogu ukazivati na potencijalne kvarove ili sajber-napade. Ovo je ilustrovano kroz analizu stvarnih događaja unutar evropske elektroenergetske mreže, gde je EIF algoritam efikasno prepoznao kritične anomalije koje su prethodile ozbiljnim incidentima.

U praktičnom smislu, implementacija EIF algoritma u alate za nadzor i dijagnostiku može značajno doprineti proaktivnom održavanju stabilnosti elektroenergetskog sistema. Ovo ne samo da poboljšava sigurnost sistema, već i optimizuje operativne performanse, čime se smanjuju troškovi i potencijalni gubici usled neplaniranih ispadanja.

Na osnovu prikazanih rezultata i analiza, može se zaključiti da je Extended Isolation Forest robustan alat koji pruža značajne prednosti za kontinuirani monitoring i unapređenje pouzdanosti elektroenergetskih sistema. Dalje istraživanje i razvoj ovog pristupa mogli bi dodatno poboljšati njegovu efikasnost i primenljivost u različitim uslovima i konfiguracijama mreža, što predstavlja značajan korak napred u oblasti detekcije anomalija.

ZAHVALNICA

Autori žele da se zahvale AD Elektromreža Srbije.

7 LITERATURA

[1] Aminifar, Farrokh, et al. "Impact of WAMS malfunction on power system reliability assessment." IEEE Transactions on Smart Grid 3.3 (2012): 1302-1309.

[2] Wang, Yang, et al. "Reliability-based incremental PMU placement." IEEE Transactions on Power Systems 29.6 (2014): 2744-2752.

[3] Su, Heng-Yi, and Chih-Wen Liu. "Estimating the voltage stability margin using PMU measurements." IEEE Transactions on Power Systems 31.4 (2015): 3221-3229.

[4] Yan, Jie, Chen-Ching Liu, and Umesh Vaidya. "PMU-based monitoring of rotor angle dynamics." IEEE Transactions on Power Systems 26.4 (2011): 2125-2133.

[5] Liu, Guoping, and Vaithianathan Venkatasubramanian. "Oscillation monitoring from ambient PMU measurements by frequency domain decomposition." 2008 IEEE International Symposium on Circuits and Systems (ISCAS). IEEE, 2008.

[6] Jiang, Quanyuan, et al. "PMU-based fault location using voltage measurements in large transmission networks." IEEE transactions on power delivery 27.3 (2012): 1644-1652.

[7] Bečejac, Vladimir B. Optimalna postavka sinhrofazorskih uređaja za obezbeđenje potpune topološke opservabilnosti primenom metode Grebnerove baze. Diss. University of Belgrade (Serbia), 2020.

[8] Martinelli, Marco, et al. "Electric power system anomaly detection using neural networks." International conference on knowledge-based and intelligent information and engineering systems. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004.

[9] Nishiya, K., J. Hasegawa, and T. Koike. "Dynamic state estimation including anomaly detection and identification for power systems." IEE proceedings C (generation, transmission and distribution). Vol. 129. No. 5. IEE, 1982.

[10] Sun, Mucun, and Jie Zhang. "Data-driven anomaly detection in modern power systems." Security of cyber-physical systems: Vulnerability and Impact (2020): 131-143.

[11] Liu, Fei Tony, Kai Ming Ting, and Zhi-Hua Zhou. "Isolation forest." 2008 eighth ieee international conference on data mining. IEEE, 2008.

[12] Hariri, Sahand, Matias Carrasco Kind, and Robert J. Brunner. "Extended isolation forest." IEEE transactions on knowledge and data engineering 33.4 (2019): 1479-1489.

[13] Hariri, S.; Carrasco Kind, M.; et al. Extended Isolation Forest. IEEE Transactions on Knowledge and Data Engineering, 2019: p. 1–1, ISSN 2326-3865, doi:10.1109/tkde.2019.2947676. Available from: http://dx.doi.org/10.1109/TKDE.2019.2947676

[14] https://www.entsoe.eu/news/2021/01/26/system-separation-in-the-continental-europe-synchronous-area-on-8-january-2021-2nd-update/

[15] https://www.entsoe.eu/news/2025/02/25/entso-e-publishes-the-final-report-on-the-grid-incident-in-south-east-europe/