

Raport Științific și Tehnic

SAGE: O Simbioză între Verificarea Satisfiabilității, Rețele Neuronale pe Grafuri și Calculul Simbolic

Conform cerințelor publicate pe site-ul UEFISCDI (<https://uefiscdi.gov.ro/proiecte-de-cercetare-pentru-stimularea-tinerelor-echipe-independente>),

prezentul raport trebuie să îndeplinească următoarele cerințe:

1. să aibă max. 20 pag;
2. să conțină descrierea științifică cu punerea în evidență a rezultatelor etapei anuale și gradul de realizare a obiectivelor;
3. să prezinte un sumar al progresului (livrabile realizate, indicatori de rezultat, diseminarea rezultatelor, justificare diferențe, dacă e cazul);
4. Să includă un rezumat executiv al activităților realizate în perioada de implementare (max. 1 pag.). Acesta poate fi publicat de către Autoritatea Contractantă în pagina web a competiției.

Itemii 1-3 sunt abordați în descrierea aferentă raportului științific și tehnic, iar itemul 4 este realizat în Anexa 1 fiind un document de sine stătător.

Identificare

Contractor: Universitatea de Vest din Timișoara (UVT)

Director de proiect: Mădălina Erașcu

Cod: PN-III-P1-1.1-TE-2021-0676

Etapa 2: Ianuarie - Decembrie 2023

Pagina web: <https://merascu.github.io/links/SAGE.html>

Rezumatul etapei

Proiectul *SAGE* (*A Symbiosis of Satisfiability Checking, Graph Neural Networks and Symbolic Computation*) abordează problema dificilă a analizei programelor prin dezvoltarea unor tehnici matematice riguroase care se ocupă de părțile logic complexe ale software-ului. În acest scop, obiectivul acestuia este să utilizeze, să rafineze și să combine metode din verificarea satisfiabilității, rețele neuronale pe grafuri și calcul simbolic pentru a face posibilă analiza software-ului care depășește capacitățile metodelor existente, deoarece tehnici avansate de calcul simbolic (teoria invarianților) și rețele neuronale pe grafuri combinate cu verificarea satisfiabilității (unelte SAT/SMT) nu sunt exploatate în metodele și instrumentele actuale de analiză a programelor.

Obiectivele etapei 2 (ianuarie - decembrie 2023) au fost definitivarea activităților din Etapa 1 (obiectivele O1 și O2) precum și dezvoltarea unei teorii generice și algoritmi pentru ruperea simetriilor și învățarea de șabloane pe baza studiilor de caz investigate în Etapa 1 (obiectiv O3).

- *Obiectiv O1*: Identificarea de simetrii și metodele de rupere a simetriilor, precum și de a învăța șabloane de probleme pentru două cazuri de utilizare relevante: managementului resurselor Cloud și rețele neuronale binarizate - varianta finală
- *Obiectiv O2*: Formularea problemei de managementului resurselor Cloud ca o problemă de rețele neuronale pe grafuri pentru a învăța șabloane de probleme pentru care găsirea optimului este mai ușoară, deoarece poate fi prezis în prealabil - varianta finală
- *Obiectiv O3*: Dezvoltarea unei teorii generice și algoritmi pentru ruperea simetriilor și învățarea de șabloane pe baza studiilor de caz investigate în Etapa 1.

Rezultatele obținute în cadrul proiectului în anul 2023 sunt:

- Evaluarea comparativă a diferite metode și formalisme de rupere a simetriilor pentru probleme de optimizare provenind de la probleme de deployment al aplicațiilor în Cloud.
- Dezvoltarea unui prototip software care permite deploymentul de aplicații în clustere de tip Kubernetes. Prototipul software utilizează tehnici de rupere a simetriilor dezvoltate de noi la punctul anterior.
- Proiectarea de arhitecturi pentru rețelele neuronale binarizate pentru recunoașterea semnelor de circulație precum și analiza robusteții acestora.
- 4 articole științifice publicate sau în curs de publicare în volumele unor conferințe/workshopuri internaționale:
 - Architecturing Binarized Neural Networks for Traffic Sign Recognition, Andreea Postovan and Mădălina Erașcu, 32nd International Conference on Artificial Neural Networks (ICANN 2023) (Springer LNCS Proceedings), <https://link.springer.com/book/10.1007/978-3-031-44207-0>
 - SAGE - A Tool for Optimal Deployments in Kubernetes Clusters, Vlad Luca and Mădălina Erașcu, The 14th IEEE International Conference on Cloud Computing Technology and Science (IEEE CloudCom), December 4-6 2023 (IEEE Proceedings, indexed ISI Web of Science).
 - Benchmarking Optimization Solvers and Symmetry Breakers for the Automated Deployment of Component-based Applications in the Cloud, Bogdan David and Mădălina Erașcu, Satisfiability Checking and Symbolic Computation 2022, Haifa, Israel, August 12, 2022 (published on CEUR Workshop Proceedings on Aug 19, 2023), affiliated with IJCAR and FLOC 2022
 - Benchmarking Local Robustness of High-Accuracy Binary Neural Networks for Enhanced Traffic Sign Recognition, Andreea Postovan and Mădălina Erașcu, Working Formal Methods 2023, Bucharest, Romania, September 21-22

La activitățile de diseminare de mai sus, se adaugă mai multe activități în sprijinul comunităților de cercetare din care facem parte (recenzii conferințe, organizare și participare la evenimente științifice, etc.) care sunt prezentate în secțiunea următoare.

Obiectivele planificate în anul 2023, cât și activitățile aferente acestora au fost realizate în totalitate, și desfășurate conform cu planul de realizare al proiectului. Prin activitățile de diseminare am contribuit la criteriile de performanță asumate în implementarea proiectului.

Descriere științifică și tehnică

Obiectivul general proiectului este de a avansa teoria de ultimă generație și algoritmi de verificare a satisfiabilității, calcul simbolic și rețele neuronale binarizate pentru rezolvarea problemelor de satisfacere a constrângerilor pe scară largă. Ne propunem să:

(O1) oferim o explicație a simetriilor (O1.1) și asemănarilor (O1.2) pentru probleme practice importante, cum ar fi (a) gestionarea resurselor în cloud, (b) verificarea rețelelor neuronale binarizate (O2) dezvoltăm metode (O2.1) pentru spargerea simetriilor în, și (O2.2) pentru învățarea șabloanelor din problemele enumerate mai sus.

(O3) inventăm teorie și algoritmi pentru (O3.1) abstractizarea simetriilor acestor studii de caz și pentru descrierea tehnicilor de rupere a simetriei folosind teoria grupurilor invariante și rezolvarea SAT/SMT; (O3.2) învățarea șabloanelor de probleme prin formalizarea problemei ca rețea neuronală pe grafuri și aplicarea bibliotecilor existente de rețele neuronale pe grafuri pentru diferite predicții

(O4) studiem eficiența computațională a tehnicilor de rupere a simetriei și a șabloanelor nou dezvoltate;

(O5) pentru, în cele din urmă, rezolvarea automată a problemelor de satisfacție a constrângerilor la scară largă.

Rezultatele estimate ale proiectului sunt:

1. rapoarte științifice și tehnice care conțin metodele originale din sfera logicii computaționale (SAT/SMT solving), algebră computațională (teoria invarianților grupurilor finite) și învățare automată (rețele neuronale pe grafuri) care combinate să permită rezolvarea automată a problemelor de satisfacție a constrângerilor la scară largă.
2. publicații științifice pentru diseminarea rezultatelor științifice obținute;
3. prototipuri software (incluse în sistemul SAGE <https://github.com/SAGE-Project>) care implementează metodele menționate la punctul 2.

Obiectivele etapei 2 (ianuarie - decembrie 2023) au fost definitivarea activităților din Etapa 1 (obiectivele O1 și O2) precum și dezvoltarea unei teorii generice și algoritmi pentru ruperea simetriilor și învățarea de șabloane pe baza studiilor de caz investigate în Etapa 1 (obiectiv O3).

- Obiectiv O1: Identificarea de simetrii și metodele de rupere a simetriilor, precum și de a învăța șabloane de probleme pentru două cazuri de utilizare relevante: managementului resurselor Cloud și rețele neuronale binarizate - varianta finală
- Obiectiv O2: Formularea problemei de managementului resurselor Cloud ca o problemă de rețele neuronale pe grafuri pentru a învăța șabloane de probleme pentru care găsirea optimului este mai ușor de găsit, deoarece poate fi prezis în prealabil - varianta finală
- Obiectiv O3: Dezvoltarea unei teorii generice și algoritmi pentru ruperea simetriilor și învățarea de șabloane pe baza studiilor de caz investigate în Etapa 1.

Activitățile științifice și tehnice aferente îndeplinirii obiectivelor sunt:

Activitate 2.1	Identificarea simetriilor și investigarea tehnicilor de rupere a simetriei pentru studiul de caz: management al resurselor în Cloud (varianta finală)
Activitate 2.2	Identificarea simetriilor și investigarea tehnicilor de rupere a simetriei pentru studiul de caz: verificarea proprietăților rețelelor neuronale binarizate (varianta finală)
Activitate 2.3	Identificarea asemănărilor și investigarea modului în care șabloanele de învățare sunt aplicabile pentru studiul de caz: managementul resurselor în Cloud (varianta preliminară)
Activitate 2.4	Tehnici de abstractizare adecvate pentru simetrii și rupere a simetriilor
Activitate 2.5	Cadru teoretic pentru tehnici de rupere a simetriei
Activitate 2.6	Cadru algoritmic pentru învățare de șabloane

Gradul de îndeplinire a obiectivelor etapei

Obiectivele științifice și tehnice ale etapei a fost îndeplinit în totalitate, cum vom prezenta mai jos. Reamintim pachetele de lucru și activitățile stabilite în descrierea inițială a proiectului. După cum se observă din tabelul de mai jos, activitățile din WP1 și WP2 sunt preponderent teoretice iar cele din WP3 le implementează, testează și rafinează. La momentul actual activitățile din WP1 sunt încheiate iar cele din WP2 și WP3 sunt în derulare, exceptând T3.1 prin care am stabilit arhitectura aplicației și toolurile folosite.

WP / Task	Name	Start Month	End Month
WP1	Case Studies Analysis		
T1.1	Identification of symmetries and investigation of the symmetry breaking	M01	M12

	techniques for the Case Study A (resource provisioning in the Cloud)		
T1.2	Identification of symmetries and investigation of the symmetry breaking techniques for the Case Study B (property verification of BNNs)	M01	M12
T1.3	Identification of similarities and investigation of how the learning templates are applicable for Case Study A (resource provisioning in the Cloud)	M01	M12
WP2	Theoretical and Algorithmic Framework		
T2.1	Suitable abstraction techniques for symmetries and symmetry breaking	M03	M18
T2.2	Theoretical framework for symmetry breaking	M03	M18
T2.3	Algorithmic framework for template learning	M03	M18
WP3	Integrated Framework		
T3.1	Architecture definition and support tools	M06	M12
T3.2	Implementation of the symmetry breaking framework	M09	M21
T3.3	Integration of the learning templates	M09	M21
T3.4	Validation of the implementation	M12	M24
WP4	Dissemination and Project Management		
T4.1	Dissemination, communication, exploitation	M01	M24
T4.2	Project Management	M01	M24

Considerăm că obiectivele tehnice ale etapei, inclusiv diseminarea, sunt îndeplinite după cum reiese din livrabilele [L1] - [L4].

- [L1] Benchmarking Optimization Solvers and Symmetry Breakers for the Automated Deployment of Component-based Applications in the Cloud, Bogdan David and Mădălina Erașcu, Satisfiability Checking and Symbolic Computation 2022, Haifa, Israel, August 12, 2022 (published on CEUR Workshop Proceedings on Aug 19, 2023), affiliated with IJCAR and FLOC 2022 - în cadrul activităților 2.1, 2.4, 2.5.
- [L2] SAGE - A Tool for Optimal Deployments in Kubernetes Clusters, Vlad Luca and Mădălina Erașcu, The 14th IEEE International Conference on Cloud Computing Technology and Science (IEEE CloudCom), December 4-6 2023 (IEEE Proceedings, indexed ISI Web of Science) - în cadrul activităților 2.1, 2.4, 2.5.
- [L3] Architecturing Binarized Neural Networks for Traffic Sign Recognition, Andreea Postovan and Mădălina Erașcu, 32nd International Conference on Artificial Neural Networks (ICANN 2023) (Springer LNCS Proceedings), <https://link.springer.com/book/10.1007/978-3-031-44207-0> - în cadrul activității 2.2.
- [L4] Benchmarking Local Robustness of High-Accuracy Binary Neural Networks for Enhanced Traffic Sign Recognition, Andreea Postovan and Mădălina Erașcu, Working Formal Methods 2023, Bucharest, Romania, September 21-22 - în cadrul activității 2.2.

La aceste livrabile se adaugă în imediata perioadă:

- Un articol intitulat *Fast and Exact Synthesis of Optimal Deployment Plans using SMT Solving and Learned Predictions* pe care intenționăm să-l trimitem la o conferință cotate ISI - în cadrul activităților 2.3, 2.6.

Diseminare

Diseminarea în cadrul *conferințelor* și *workshopurilor* s-a făcut prin [L1] - [L4]. Articolele au fost însoțite de prezentări.

Mădălina Erașcu a participat la *Seminarul Dagstuhl 23401 Automated mathematics: integrating proofs, algorithms and data*, desfășurat în perioada 1-6 octombrie 2023, la Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik din Germania. Prezentarea pe care a susținut-o și interacțiunea cu ceilalți participanți s-au axat pe înțelegerea drumului de la simetrii la spărgătoarele de simetrie ale

problemelor de optimizare a constrângerilor care apar în implementarea în Cloud a aplicațiilor bazate pe componente.

În luna iulie, studenții din cadrul proiectului și-au susținut dizertațiile într-un mod excelent:

- Marcus Ilisie. *A Symbiosis of Constraint Optimization, Symmetries and Symmetry Breaking for Scalable Cloud Deployment Problems*.
- Eduard Laitin. *Speeding Up the Deployment in the Cloud of Component-based Applications using Graph Neural Networks and SMT solving*.
- Vlad Luca. *SAGE - A Tool for Optimal Deployments in Kubernetes Clusters*.
- Andreea Postovan. *Binarized Neural Networks for Traffic Sign Recognition: Training and Verification*.

Alte Activități

- Vlad Luca este membru în Comitetul de Evaluare a Artefactelor la cea de-a 30-a Conferință Internațională privind Instrumente și Algoritmi pentru Construcția și Analiza Sistemelor (TACAS) 2024, care va avea loc în perioada 6-13 aprilie 2024, în Luxembourg Ville, Luxembourg (<https://tacas.info/artifacts-24.php>).
- Mădălina Erașcu este membru al Comitetului de Program pentru 19th European Dependable Computing Conference, Student Forum, desfășurată la KU Leuven în perioada 8-11 aprilie 2024, în Leuven, Belgia (<https://edcc2024.esat.kuleuven.be/studentforum.html>).
- Mădălina Erașcu este subrecenzor pentru The 14th International Symposium on Frontiers of Combining Systems (FroCoS 2023), Institutul de Informatică, Robotică și Cibernetică (CIIRC) al Universității Tehnice din Praga (CTU), Republica Cehă, în perioada 20-22 septembrie 2023.
- Benchmarkurile propuse în [L4], pentru evaluarea robusteții locale a modelelor de rețele neuronale binarizate, au fost incluse în competiția 4th International Verification of Neural Networks Competition (VNN-COMP'23). Modelele antrenate implică straturi precum convoluție binarizată (furnizată de biblioteca Larq), max pooling, normalizare în lot și densitate (https://github.com/ChristopherBrix/vnncomp2023_benchmarks/tree/main/benchmarks/traffic_signs_recognition).
- Mădălina Erașcu este membru al Comitetului de Program pentru cea de-a 25-a Simpozion Internațional privind Algoritmi Simbolici și Numerici pentru Calculul Științific (SYNASC 2023), desfășurat la LORIA, Nancy, Franța, pentru secțiunile de Programare Logică și Calcul Simbolic (<https://synasc.ro/2023/logic-and-programming/>), precum și la Workshopul privind Teoria Contractelor Inteligente și Aplicații (TOSCA) (<https://synasc.ro/2023/tosca2023/>).
- Mădălina Erașcu este membru al Comitetului de Program pentru 32nd International Conference on Artificial Neural Networks (ICANN 2023), organizată de School of Engineering of Democritus University of Thrace, Greece (<https://e-nns.org/icann2023/organization/programme-committee/>).
- Mădălina Erașcu a susținut o prezentare la The Tenth Congress of Romanian Mathematicians desfășurat în perioada 30 iunie - 5 iulie 2023, la Pitești, România, în cadrul Sesiunii Speciale de Logică și Aplicații. Congresul este organizat în comun de Secția de Științe Matematice a Academiei Române, Societatea Matematică Română, Institutul de Matematică Simion Stoilow al Academiei Române, Facultatea de Matematică și Informatică a Universității din București și Universitatea din Pitești (<http://www.imar.ro/~congmatro10/ProgramCRM10.pdf>).
- În luna februarie 2023 am găzduit întâlnirea Grupului de Lucru 3 (Verificarea Programelor) a proiectului COST EuroProofNet (<https://europroofnet.github.io/wg3-timisoara/>).
- Mădălina Erașcu este membru al Comitetului de Program pentru cea de-a 30-a Conferință Internațională privind Instrumente și Algoritmi pentru Construcția și Analiza Sistemelor

(TACAS) 2024, care va avea loc în perioada 6-13 aprilie 2024, în Luxembourg Ville, Luxembourg (<https://etaps.org/2024/conferences/tacas/>).

Buget și echipă

Costurile de personal sunt conforme cu efortul depus și estimările inițiale, cu observația că echipa a suferit unele modificări prin absolvirea masterului de către studenții implicați în proiect. Dintre aceștia, studentul Vlad Luca a devenit doctorand în cadrul Școlii Doctorale de Informatică de la 1 octombrie. Am cooptat în echipa și pe studenta Anda Leșeanu, masterandă în anul 1 la specializarea Artificial Intelligence and Distributed Computing (master de cercetare în cadrul Departamentului de Informatică, Universitatea de Vest din Timișoara).

Concluzii

- Obiectivele etapei au fost îndeplinite în totalitate.
- Implementarea conceptelor este în faza avansată.
- Au existat activități de diseminare variate.
- Informații actualizate despre proiect se pot afla la <https://merascu.github.io/links/SAGE.html>.
- Prototipurile și benchmarkurile aferente proiectului se află la <https://github.com/SAGE-Project>.

Anexa 1

Rezumat executiv al activităților realizate în cadrul proiectului *SAGE: O Simbioză între Verificarea Satisfiabilității, Rețele Neuronale pe Grafuri și Calculul Simbolic*

Contractor: Universitatea de Vest din Timișoara (UVT)

Cod: PN-III-P1-1.1-TE-2021-0676

Etapa 2: Ianuarie - Decembrie 2023

Pagina web: <https://merascu.github.io/links/SAGE.html>

Problemă

Viața noastră depinde de software și sisteme care folosesc software. În zilele noastre, software-ul este complex și, prin urmare, este predispus la erori. Pentru a depăși această problemă, avem nevoie de metode pentru a sprijini dezvoltarea software-ului.

Soluție

Acest proiect abordează problema provocatoare a analizei software prin dezvoltarea de tehnici matematice riguroase care se ocupă de părțile complexe ale software-ului. În acest scop, vom folosi, rafina și combina metode de verificare a satisfacției de constrângeri, rețele neuronale pe grafuri și calcul simbolic. Acestea vor face posibilă analiza software-ului care depășește puterea metodelor existente, deoarece tehnici avansate de calcul simbolic (teoria invarianților), rețele neuronale pe grafuri și combinarea lor cu metode de satisfacere de constrângeri (unelte SAT/SMT) nu este exploatată în literatură.

Metodologie

Vom dezvolta noi teorii și algoritmi care detectează simetrii ale problemei și similitudini între probleme din aceeași familie prin combinarea teoriei invarianților, a rețelelor neuronale pe grafuri și a metodelor SAT/SMT care, spre deosebire de metodele curente, ar permite rezolvarea eficientă a problemelor intractabile din punct de vedere teoretic.

Proiectul nostru de cercetare vizează analiza software-ului complex și urmărește trei direcții principale de cercetare: (1) Metode de eliminare a simetriilor și de învățare a similitudinilor între probleme, (2) Teorie și algoritmi, (3) Instrumente și evaluări.

Rezultatele obținute în cadrul etapei 2 sunt:

1. Evaluarea comparativă a diferite metode și formalisme de rupere a simetriilor pentru probleme de optimizare provenind de la probleme de deployment al aplicațiilor în Cloud.
2. Dezvoltarea unui prototip software care permite deploymentul de aplicații în clustere de tip Kubernetes. Prototipul software utilizează tehnici de rupere a simetriilor dezvoltate de noi la punctul anterior.
3. Proiectarea de arhitecturi pentru rețele neuronale binarizate pentru recunoașterea semnelor de circulație precum și analiza robusteții acestora.
4. 4 articole științifice publicate sau în curs de publicare în volumele unor conferințe/workshopuri internaționale.

La activitățile de diseminare de la punctul 4, se adaugă mai multe activități în sprijinul comunităților de cercetare (recenzii conferințe, organizare și participare la evenimente științifice) care sunt prezentate exhaustiv în secțiunea următoare.

Obiectivele planificate pe anul 2023, cât și activitățile aferente acestora au fost realizate în totalitate, și desfășurate conform cu planul de realizare al proiectului. Prin activitățile de diseminare am contribuit la criteriile de performanță asumate în implementarea proiectului.

Concluziile Etapei 2

- Obiectivele etapei au fost îndeplinite în totalitate.
- Implementarea conceptelor este în faza avansată.
- Au existat activități de diseminare variate.
- Informații actualizate despre proiect se pot afla la <https://merascu.github.io/links/SAGE.html>.
- Prototipurile și benchmarkurile aferente proiectului se află la <https://github.com/SAGE-Project>.