

ΠΡΟΤΕΙΝΟΜΕΝΑ ΘΕΜΑΤΑ ΔΙΠΛΩΜΑΤΙΚΩΝ ΕΡΓΑΣΙΩΝ ΓΙΑ ΤΟ ΧΕΙΜΕΡΙΝΟ ΕΞΑΜΗΝΟ 2020-2021

Εισηγητής: Νίκος Πλόσκας (nploskas@uowm.gr), Επίκουρος Καθηγητής ΤΗΜΜΥ

1. Αυτοματοποιημένη εύρεση βέλτιστων παραμέτρων λογισμικού (Auto-tuning software parameters)

Περιγραφή: Τα συστήματα λογισμικού για επιστημονικούς υπολογισμούς διαθέτουν διάφορες παραμέτρους για την ταχύτερη εκτέλεσή τους ή/και για την εύρεση καλύτερων λύσεων. Οι δημιουργοί αυτών των λογισμικών έχουν διεξάγει πολλά πειράματα για την εύρεση κατάλληλων προκαθορισμένων τιμών για αυτές τις παραμέτρους. Ωστόσο, τα συστήματα λογισμικού δεν έχουν πάντα την επιθυμητή απόδοση σε ορισμένες κλάσεις προβλημάτων. Στόχος της εργασίας αυτής είναι η υλοποίηση και υπολογιστική σύγκριση αλγορίθμων για την εύρεση βέλτιστων παραμέτρων λογισμικού.

Ενδεικτική βιβλιογραφία:

1. Arcuri, A., & Fraser, G. (2013). Parameter tuning or default values? An empirical investigation in search-based software engineering. *Empirical Software Engineering*, 18(3), 594-623.
2. Elkhodary, A., Esfahani, N., & Malek, S. (2010). FUSION: a framework for engineering self-tuning self-adaptive software systems. In *Proceedings of the eighteenth ACM SIGSOFT international symposium on Foundations of software engineering* (pp. 7-16). ACM.
3. Katagiri, T., Kise, K., Honda, H., & Yuba, T. (2003). Fiber: A generalized framework for auto-tuning software. In *International Symposium on High Performance Computing* (pp. 146-159). Springer, Berlin, Heidelberg.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Προγραμματισμός, Αριθμητική Ανάλυση.

2. Υλοποίηση και υπολογιστική σύγκριση βελτιωτικών ευρετικών μεθόδων για το πρόβλημα του πλανόδιου πωλητή (Implementation and computational comparison of improvement heuristics for the traveling salesman problem)

Περιγραφή: Το πρόβλημα του πλανόδιου πωλητή είναι ένα κλασικό πρόβλημα συνδυαστικής βελτιστοποίησης. Υπάρχουν πολλές ευρετικές διαδικασίες που δέχονται ως είσοδο μια λύση του προβλήματος και τη βελτιώνουν. Στόχος της εργασίας είναι η μελέτη, η υλοποίηση και η υπολογιστική σύγκριση βελτιωτικών ευρετικών μεθόδων για την επίλυση του προβλήματος ανάθεσης πόρων.

Ενδεικτική βιβλιογραφία:

4. Helsgaun, K. (2009). General k-opt submoves for the Lin–Kernighan TSP heuristic. *Mathematical Programming Computation*, 1(2-3), 119-163.
5. Karapetyan, D., & Gutin, G. (2011). Lin–Kernighan heuristic adaptations for the generalized traveling salesman problem. *European Journal of Operational Research*, 208(3), 221-232.
6. Laporte, G. (2010). A concise guide to the traveling salesman problem. *Journal of the Operational Research Society*, 61(1), 35-40.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Προγραμματισμός.

3. Σύγκριση μεθόδων μείωσης διαστάσεων των δεδομένων (Comparison of dimension reduction techniques)

Περιγραφή: Πολλές εφαρμογές χρησιμοποιούν διάφορους αισθητήρες για την συλλογή δεδομένων. Τα δεδομένα που συλλέγονται είναι μεγάλα και ανομοιόμορφα, οπότε είναι δύσκολη η επεξεργασία τους και η λήψη αποφάσεων με αυτά. Μια τεχνική που χρησιμοποιείται πριν εφαρμοστεί κάποιος αλγόριθμος μηχανικής μάθησης είναι η μείωση των διαστάσεων των δεδομένων (dimension reduction). Στόχος της εργασίας είναι η σύγκριση τεχνικών μείωσης των διαστάσεων δεδομένων πριν την εκτέλεση αλγορίθμων μηχανικής μάθησης.

Ενδεικτική βιβλιογραφία:

1. Papageorgiou, A., Cheng, B., & Kovacs, E. (2015). Real-time data reduction at the network edge of Internet-of-Things systems. In: *11th International Conference on Network and Service Management (CNSM)*, 284-291.
2. Li, L. (2010). Dimension Reduction for High-Dimensional Data. In: Bang, H., Zhou, X., van Epps, H., & Mazumdar, M. (eds.) *Statistical Methods in Molecular Biology. Methods in Molecular Biology (Methods and Protocols)*, 620, Humana Press, Totowa, NJ.
3. Van Der Maaten, L., Postma, E., & Van den Herik, J. (2009). Dimensionality reduction: a comparative. *Journal of Machine Learning Research*, 10(66-71), 13.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Τεχνητή Νοημοσύνη, Προγραμματισμός, Αριθμητική Ανάλυση, Στατιστική.

4. Σύγκριση αλγορίθμων δημιουργίας υποκατάστατων μοντέλων (Comparison of surrogate-based methods)

Περιγραφή: Σε αρκετά προβλήματα βελτιστοποίησης δεν υπάρχει η αναλυτική μορφή της αντικειμενικής συνάρτησης, οπότε χειριζόμαστε το πρόβλημα ως ένα πρόβλημα μαύρου κουτιού (black box optimization). Υπάρχουν διάφοροι αλγόριθμοι που δημιουργούν υποκατάστατα μοντέλα (surrogate models) για να προσεγγίσουν τις αντικειμενικές συναρτήσεις σε αυτά τα προβλήματα. Στόχος της εργασίας είναι η σύγκριση αλγορίθμων δημιουργίας υποκατάστατων μοντέλων.

Ενδεικτική βιβλιογραφία:

1. Owen, N. E., Challenor, P., Menon, P. P., & Bennani, S. (2017). Comparison of surrogate-based uncertainty quantification methods for computationally expensive simulators. *SIAM/ASA Journal on Uncertainty Quantification*, 5(1), 403-435.
2. Qian, Z., Seepersad, C. C., Joseph, V. R., Allen, J. K., & Wu, C. J. (2006). Building surrogate models based on detailed and approximate simulations. *Journal of Mechanical Design*, 128(4), 668-677.
3. Tabatabaei, M., Hakanen, J., Hartikainen, M., Miettinen, K., & Sindhya, K. (2015). A survey on handling computationally expensive multiobjective optimization problems using surrogates: non-nature inspired methods. *Structural and Multidisciplinary Optimization*, 52(1), 1-25.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Τεχνητή Νοημοσύνη, Προγραμματισμός, Αριθμητική Ανάλυση, Στατιστική.

5. Υλοποίηση αλγορίθμων γραμμικού προγραμματισμού (Implementation of linear programming algorithms)

Περιγραφή: Τα προβλήματα γραμμικού προγραμματισμού εμφανίζονται σε διάφορες εφαρμογές και η βελτίωση των ήδη υπαρχόντων αλγορίθμων είναι μεγάλης σημασίας. Στόχος της εργασίας είναι η μελέτη και η υπολογιστική σύγκριση αλγορίθμων γραμμικού προγραμματισμού τύπου simplex. Θα υλοποιηθούν και θα συγκριθούν διάφορες μέθοδοι σε κάθε βήμα των αλγορίθμων τύπου simplex.

Ενδεικτική βιβλιογραφία:

1. Dantzig, G. B. (1998). *Linear programming and extensions* (Vol. 48). Princeton University Press.
2. Vanderbei, R. J. (2015). *Linear programming*. Heidelberg: Springer.
3. Ploskas, N., & Samaras, N. (2017). *Linear programming using MATLAB®* (Vol. 127). Switzerland: Springer.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Προγραμματισμός, Επιχειρησιακή Έρευνα, Αριθμητική Ανάλυση.

6. Σύγκριση παράλληλων αλγορίθμων ταξινόμησης (Comparison of parallel sorting algorithms)

Περιγραφή: Το πρόβλημα ταξινόμησης είναι από τα πιο απλά προβλήματα της πληροφορικής που πολλοί αλγόριθμοι χρειάζονται να λύνουν πολλές φορές. Υπάρχουν διάφοροι παράλληλοι αλγόριθμοι ταξινόμησης. Στόχος της εργασίας είναι η σύγκριση παράλληλων αλγορίθμων ταξινόμησης.

Ενδεικτική βιβλιογραφία:

1. Inoue, H., Moriyama, T., Komatsu, H., & Nakatani, T. (2007). AA-sort: A new parallel sorting algorithm for multi-core SIMD processors. In *IEEE 16th International Conference on Parallel Architecture and Compilation Techniques (PACT 2007)*, 189-198.
2. Herruzo, E., Ruiz, G., Benavides, J. I., & Plata, O. (2007). A new parallel sorting algorithm based on odd-even mergesort. In *IEEE 15th EUROMICRO International Conference on Parallel, Distributed and Network-Based Processing (PDP'07)*, 18-22.
3. Akl, S. G. (2014). *Parallel sorting algorithms* (Vol. 12). Academic press.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Προγραμματισμός.

7. Σύγκριση αλγορίθμων ταξινόμησης αλφαριθμητικών (Comparison of string sorting algorithms)

Περιγραφή: Το πρόβλημα ταξινόμησης είναι από τα πιο απλά προβλήματα της πληροφορικής που πολλοί αλγόριθμοι χρειάζονται να λύνουν πολλές φορές. Υπάρχουν διάφοροι αλγόριθμοι ταξινόμησης αλφαριθμητικών. Στόχος της εργασίας είναι η σύγκριση αλγορίθμων ταξινόμησης αλφαριθμητικών.

Ενδεικτική βιβλιογραφία:

1. Bentley, J. L., & Sedgewick, R. (1997, January). Fast algorithms for sorting and searching strings. In *Proceedings of the eighth annual ACM-SIAM symposium on Discrete algorithms*, 360-369.
2. Neelima, B., Narayan, A. S., & Prabhu, R. G. (2014, December). String sorting on multi and many-threaded architectures: A comparative study. In *IEEE 2014 International Conference on High Performance Computing and Applications (ICHPCA)*, 1-6.
3. Angrish, R., & Garg, D. (2011). Efficient string sorting algorithms: Cache-aware and cache-oblivious. *International Journal of Soft Computing and Engineering*, 1(2), 12-16.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Προγραμματισμός.

8. Αλγόριθμοι μηχανικής μάθησης για τον εντοπισμό ανωμαλιών (προβλημάτων/επιθέσεων) σε κινητές συσκευές (Machine learning algorithms for anomaly detection in mobile devices)

Περιγραφή: Τα δεδομένα που προέρχονται από τις μετρήσεις κίνησης δικτύου μιας κινητής συσκευής μπορούν να χρησιμοποιηθούν για τον εντοπισμό ανωμαλιών και επιθέσεων ασφάλειας. Στόχος της εργασίας είναι η εφαρμογή και αξιολόγηση αλγορίθμων μηχανικής μάθησης για τον έγκαιρο εντοπισμό ανωμαλιών και επιθέσεων σε κινητές συσκευές.

Ενδεικτική βιβλιογραφία:

1. Ye, Y., Li, T., Adjero, D., & Iyengar, S. S. (2017). A survey on malware detection using data mining techniques. *ACM Computing Surveys*, 50(3), 1-40.
2. Majeed, K., Jing, Y., Novakovic, D., & Ouazzane, K. (2014). Behaviour based anomaly detection for smartphones using machine learning algorithm. In *Proceedings of the International conference on Computer Science and Information Systems*, 17-18.
3. Qamar, A., Karim, A., & Chang, V. (2019). Mobile malware attacks: Review, taxonomy & future directions. *Future Generation Computer Systems*, 97, 887-909.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Τεχνητή Νοημοσύνη, Προγραμματισμός, Αριθμητική Ανάλυση, Στατιστική.

9. Αλγόριθμοι μηχανικής μάθησης για τον εντοπισμό ανωμαλιών (προβλημάτων/επιθέσεων) σε εξυπηρετητές (Machine learning algorithms for anomaly detection in servers)

Περιγραφή: Τα δεδομένα που προέρχονται από τα αρχεία καταγραφής των εξυπηρετητών (είτε πρόκειται για εξυπηρετητές εφαρμογών είτε για εξυπηρετητές βάσεων δεδομένων) μπορούν να χρησιμοποιηθούν για τον εντοπισμό ανωμαλιών και επιθέσεων ασφάλειας. Στόχος της εργασίας είναι η εφαρμογή και αξιολόγηση αλγορίθμων μηχανικής μάθησης για τον έγκαιρο εντοπισμό ανωμαλιών και επιθέσεων σε εξυπηρετητές.

Ενδεικτική βιβλιογραφία:

1. Cao, Q., Qiao, Y., & Lyu, Z. (2017, December). Machine learning to detect anomalies in web log analysis. In 2017 3rd IEEE International Conference on Computer and Communications (ICCC), 519-523.
2. Juvonen, A., Sipola, T., & Hämmäläinen, T. (2015). Online anomaly detection using dimensionality reduction techniques for HTTP log analysis. *Computer Networks*, 91, 46-56.
3. Du, M., Li, F., Zheng, G., & Srikumar, V. (2017, October). Deeplog: Anomaly detection and diagnosis from system logs through deep learning. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 1285-1298.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Τεχνητή Νοημοσύνη, Προγραμματισμός, Αριθμητική Ανάλυση, Στατιστική.

10. Αλγόριθμοι ενισχυτικής μάθησης για την επίλυση του προβλήματος δρομολόγησης οχημάτων (Reinforcement learning algorithms for the vehicle routing problem)

Περιγραφή: Το πρόβλημα δρομολόγησης οχημάτων είναι ένα κλασικό πρόβλημα συνδυαστικής βελτιστοποίησης. Η ενισχυτική μάθηση μπορεί να χρησιμοποιηθεί σε περιβάλλοντα που οι πληροφορίες για το περιβάλλον (π.χ. απόσταση μεταξύ δύο κόμβων) δεν είναι γνωστές. Στόχος της εργασίας είναι η υλοποίηση και αξιολόγηση αλγορίθμων ενισχυτικής μάθησης για το πρόβλημα δρομολόγησης οχημάτων.

Ενδεικτική βιβλιογραφία:

1. Nazari, M., Oroojlooy, A., Snyder, L., & Takác, M. (2018). Reinforcement learning for solving the vehicle routing problem. In *Advances in Neural Information Processing Systems*, 9839-9849.
2. Silva, M. A. L., de Souza, S. R., Souza, M. J. F., & Bazzan, A. L. C. (2019). A reinforcement learning-based multi-agent framework applied for solving routing and scheduling problems. *Expert Systems with Applications*, 131, 148-171.
3. Zhao, J., Mao, M., Zhao, X., & Zou, J. (2020). A Hybrid of Deep Reinforcement Learning and Local Search for the Vehicle Routing Problems. *IEEE Transactions on Intelligent Transportation Systems*.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Τεχνητή Νοημοσύνη, Προγραμματισμός, Αριθμητική Ανάλυση, Στατιστική.

11. Αλγόριθμοι ενισχυτικής μάθησης για την επίλυση του προβλήματος του πλανόδιου πωλητή (Reinforcement learning algorithms for the traveling salesman problem)

Περιγραφή: Το πρόβλημα του πλανόδιου πωλητή είναι ένα κλασικό πρόβλημα συνδυαστικής βελτιστοποίησης. Η ενισχυτική μάθηση μπορεί να χρησιμοποιηθεί σε περιβάλλοντα που οι πληροφορίες για το περιβάλλον (π.χ. απόσταση μεταξύ δύο κόμβων) δεν είναι γνωστές. Στόχος της εργασίας είναι η υλοποίηση και αξιολόγηση αλγορίθμων ενισχυτικής μάθησης για το πρόβλημα του πλανόδιου πωλητή.

Ενδεικτική βιβλιογραφία:

1. Gambardella, L. M., & Dorigo, M. (1995). Ant-Q: A reinforcement learning approach to the traveling salesman problem. In *Machine Learning Proceedings 1995*, 252-260.

2. Liu, F., & Zeng, G. (2009). Study of genetic algorithm with reinforcement learning to solve the TSP. *Expert Systems with Applications*, 36(3), 6995-7001.
3. Miki, S., Yamamoto, D., & Ebara, H. (2018, August). Applying deep learning and reinforcement learning to traveling salesman problem. In *2018 International Conference on Computing, Electronics & Communications Engineering (iCCECE)*, 65-70.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Τεχνητή Νοημοσύνη, Προγραμματισμός, Αριθμητική Ανάλυση, Στατιστική.

12. Αλγόριθμοι ενισχυτικής μάθησης για την επίλυση του προβλήματος πακετοποίησης (Reinforcement learning algorithms for the bin packing problem)

Περιγραφή: Το πρόβλημα πακετοποίησης είναι ένα κλασικό πρόβλημα συνδυαστικής βελτιστοποίησης. Η ενισχυτική μάθηση μπορεί να χρησιμοποιηθεί σε περιβάλλοντα που οι πληροφορίες για το περιβάλλον (π.χ. δεν είναι γνωστά εκ των προτέρων όλα τα κιβώτια) δεν είναι γνωστές. Στόχος της εργασίας είναι η υλοποίηση και αξιολόγηση αλγορίθμων ενισχυτικής μάθησης για το πρόβλημα πακετοποίησης.

Ενδεικτική βιβλιογραφία:

1. Hu, H., Zhang, X., Yan, X., Wang, L., & Xu, Y. (2017). Solving a new 3d bin packing problem with deep reinforcement learning method. arXiv preprint arXiv:1708.05930.
2. Verma, R., Singhal, A., Khadilkar, H., Basumatary, A., Nayak, S., Singh, H. V., ... & Sinha, R. (2020). A Generalized Reinforcement Learning Algorithm for Online 3D Bin-Packing. arXiv preprint arXiv:2007.00463.
3. Zhao, H., She, Q., Zhu, C., Yang, Y., & Xu, K. (2020). Online 3D Bin Packing with Constrained Deep Reinforcement Learning. arXiv preprint arXiv:2006.14978.

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Τεχνητή Νοημοσύνη, Προγραμματισμός, Αριθμητική Ανάλυση, Στατιστική.

13. Υλοποίηση και σύγκριση τεχνικών δειγματοληψίας για μεθόδους μηχανικής μάθησης (Implementation and comparison of sampling techniques for machine learning methods)

Περιγραφή: Σε πολλές εφαρμογές μηχανικής μάθησης θέλουμε να μάθουμε τη συμπεριφορά ενός εκτελέσιμου για το οποίο δεν έχουμε καμιά πληροφορία. Τα προβλήματα αυτά ονομάζονται προβλήματα μαύρου κουτιού. Αρχικά, χρειαζόμαστε να κάνουμε δειγματοληψία σε κάποια σημεία του χώρου για να βρούμε την τιμή που επιστρέφει το εκτελέσιμο. Οι τιμές αυτές στη συνέχεια χρησιμοποιούνται από μία μέθοδο μηχανικής μάθησης για την εκμάθηση της συμπεριφοράς του μαύρου κουτιού. Στόχος της εργασίας είναι η σύγκριση μεθόδων δειγματοληψίας πριν την εκτέλεση αλγορίθμων μηχανικής μάθησης.

Ενδεικτική βιβλιογραφία:

1. Tenne, Y. (2015). Initial sampling methods in metamodel-assisted optimization. *Engineering with Computers*, 31(4), 661-680.

2. Gong, W., Duan, Q. Y., Li, J. D., Wang, C., Di, Z. H., Ye, A. Z., ... & Dai, Y. J. (2016). An intercomparison of sampling methods for uncertainty quantification of environmental dynamic models. *J. Environ. Inf.*, 28(1), 11-24.
3. Kotti, M., Fakhfakh, M., & Tlelo-Cuautle, E. (2018, February). Effect of the design space sampling on the design performances. In *2018 IEEE 9th Latin American Symposium on Circuits & Systems (LASCAS)*, 1-4).

Απαιτούμενες γνώσεις: Αλγόριθμοι και Δομές Δεδομένων, Τεχνητή Νοημοσύνη, Προγραμματισμός, Αριθμητική Ανάλυση, Στατιστική.