

ΝΙΚΟΛΑΟΣ Α. ΜΗΤΤΑΣ

ΒΙΟΓΡΑΦΙΚΟ ΣΗΜΕΙΩΜΑ

ΘΕΣΣΑΛΟΝΙΚΗ 2012

ΠΡΟΣΩΠΙΚΑ ΣΤΟΙΧΕΙΑ

| | |
|--|--|
| <u>Όνοματεπώνυμο:</u> | Νικόλαος Μήττας |
| <u>Πατρώνυμο:</u> | Αθανάσιος |
| <u>Ημερομηνία και τόπος γέννησης:</u> | 15 Μαΐου 1981, Καβάλα |
| <u>Οικογενειακή κατάσταση:</u> | Άγαμος |
| <u>Διεύθυνση κατοικίας:</u> | Αναγνωστάκη 13, Καβάλα |
| <u>Τηλέφωνο:</u> | 2510-441732, 2510-225871 |
| <u>Κινητό Τηλέφωνο</u> | 6932879555 |
| <u>e-mail:</u> | nmittas@csd.auth.gr |

ΣΠΟΥΔΕΣ

- **1999:** Απολυτήριο Δευτεροβάθμιας Εκπαίδευσης, 3ο Γενικό Λύκειο Καβάλας.
- **2003:** Πτυχίο Τμήματος Μαθηματικών, Σχολή Θετικών Επιστημών Πανεπιστημίου Κρήτης.
- **2005:** Μεταπτυχιακό Δίπλωμα Εξειδίκευσης στην Πληροφορική στην Κατεύθυνση «Πληροφοριακά Συστήματα» από το τμήμα Πληροφορικής της Σχολής Θετικών Επιστημών του Α.Π.Θ.
- **2009:** Διδακτορικό Δίπλωμα από το τμήμα Πληροφορικής της Σχολής Θετικών Επιστημών του Α.Π.Θ.
- **2011:** Μεταδιδακτορική Έρευνα στο τμήμα Ηλεκτρολόγων Μηχανικών και Τεχνολογιών Πληροφορικής του Τεχνολογικού Πανεπιστημίου Κύπρου στο γνωστικό αντικείμενο Διαχείριση Έργων Λογισμικού –Εκτίμηση Κόστους Λογισμικού.

ΑΤΟΜΙΚΕΣ ΔΕΞΙΟΤΗΤΕΣ ΚΑΙ ΙΚΑΝΟΤΗΤΕΣ

- Αγγλικά: Proficiency in English and First Certificate in English, Cambridge.
- Γερμανικά: Zertifikat, Goethe - Institut

ΑΚΑΔΗΜΑΪΚΗ ΕΜΠΕΙΡΙΑ

- Επιστημονικός Συνεργάτης στο ΤΕΙ Καβάλας (2004 – σήμερα).
- Προσκεκλημένος Διδάσκων στο Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης Τμήμα Πληροφορικής 2009-σήμερα).
- Ωρομίσθιος Καθηγητής στο ΙΕΚ Τριανδρίας (2005-2008).

ΔΙΔΑΚΤΙΚΟ ΕΡΓΟ

Διδασκαλία των ακόλουθων μαθημάτων:

Ως Προσκεκλημένος Διδάσκων στο **Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης** (2010-σήμερα):

- ο Διατμηματικό Πρόγραμμα Μεταπτυχιακών Σπουδών (ΔΠΜΣ) στην «Πληροφορική και Διοίκηση» στα Τμήματα Πληροφορικής και Οικονομικών Επιστημών.
 - ι. *Μεθοδολογία Έρευνας & Ανάλυση Δεδομένων* (Θεωρία και Εκμάθηση του στατιστικού προγράμματος SPSS)

Ως Επιστημονικός Συνεργάτης στο **ΤΕΙ Καβάλας** (2004 – σήμερα):

| Μάθημα | Τμήματα | Ακαδημαϊκό Έτος | Σύνολο |
|---|---|------------------------|--------|
| <i>Μαθηματικά I (Απειροστικός Λογισμός I)</i> (Θεωρία και Ασκήσεις Πράξεις) | Τεχνολογίας Πετρελαίου & Φυσικού Αερίου | 2005-2010 2011-2012 | 5 έτη |
| <i>Μαθηματικά I (Απειροστικός Λογισμός I)</i> (Θεωρία και Ασκήσεις Πράξεις) | Μηχανολογίας | 2005-2006 | 1 έτος |
| <i>Μαθηματικά II (Απειροστικός Λογισμός II)</i> (Θεωρία και Ασκήσεις Πράξεις) | Τεχνολογίας Πετρελαίου & Φυσικού Αερίου | 2004-2005 2009-2012 | 4 έτη |
| <i>Μαθηματικά II (Πιθανότητες-Στατιστική)</i> (Θεωρία και Ασκήσεις Πράξεις) | Ηλεκτρολογίας | 2011-2012 | 2 έτη |
| <i>Μαθηματικά III (Σειρές Fourier-Μετασχηματισμός Laplace)</i> (Θεωρία και Ασκήσεις Πράξεις) | Ηλεκτρολογίας | 2011 | 1 έτος |
| <i>Ποσοτικές Μέθοδοι στη Διοίκηση Επιχειρήσεων</i> (Θεωρία και Ασκήσεις Πράξεις) | Διοίκηση Επιχειρήσεων | 2012 | 1 έτος |
| <i>Στατιστική Ανάλυση</i> (Εργαστήριο) | Τεχνολογίας Πετρελαίου & Φυσικού Αερίου | 2006-2009 | 3 έτη |
| <i>Εισαγωγή στη Στατιστική των Επιχειρήσεων</i> (Εργαστήριο) | Διοίκησης Επιχειρήσεων | 2007-2009 2012 | 3 έτη |
| <i>Στατιστική Επιχειρήσεων</i> (Εργαστήριο) | Λογιστικής | 2008-2009 | 1 έτος |
| <i>Πληροφορική I</i> (Εργαστήριο) | Τεχνολογίας Πετρελαίου & Φυσικού Αερίου | 2004-2007 | 3 έτη |
| <i>Εργαλεία Πληροφορικής</i> (Εργαστήριο) | Ηλεκτρολογίας | 2005-2007 2009-2010 | 3 έτη |

Ως υποψήφιος Διδάκτορας στο **Τμήμα Πληροφορικής του ΑΠΘ** (2005-2009):

- ι. *Ανώτερα Μαθηματικά II* (Διδασκαλία φροντιστηριακών ασκήσεων).

ii. *Μεθοδολογία Έρευνας & Ανάλυση Δεδομένων* (Εκμάθηση του Στατιστικού Προγράμματος SPSS-Διδασκαλία εργαστηριακών ασκήσεων).

Ως Ωρομίσθιος Καθηγητής στο **Ι.Ε.Κ. Τριανδρίας Θεσσαλονίκης** (2005 – 2008):

| Μάθημα | Εξάμηνο |
|---------------------------------|---|
| <i>Χρήση Η/Υ</i> | Χειμερινό Εξάμηνο 2005-2006 Χειμερινό Εξάμηνο 2006-2007 Χειμερινό Εξάμηνο 2007-2008 Εαρινό Εξάμηνο 2007-2008 |
| <i>Λογιστικά Φύλλα</i> | Εαρινό Εξάμηνο 2005-2006 |
| <i>Εισαγωγή στη Πληροφορική</i> | Χειμερινό Εξάμηνο 2006-2007 |
| <i>Στατιστική II</i> | Εαρινό Εξάμηνο 2006-2007 |
| <i>Βιοστατιστική I</i> | Εαρινό Εξάμηνο 2007-2008 |

ΧΡΗΜΑΤΟΔΟΤΟΥΜΕΝΑ ΕΡΓΑ

Συμμετοχή

1. «Υπηρεσίες για τη σύνδεση προπτυχιακών και μεταπτυχιακών φοιτητών με την αγορά», *Τεχνική υποστήριξη*, Φορέας: *Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης Τμήμα Πληροφορικής*, Διάστημα σύμβασης: *Απρίλιος 2006 - Ιούνιος 2006*, (Κωδικός προγράμματος 21476).
2. «Παρατηρητήριο ποιότητας λογισμικού για λογισμικό ανοικτού κώδικα», *Έρευνα στην Τεχνολογία Λογισμικού*, Φορέας: *Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης Τμήμα Πληροφορικής*, Διάστημα σύμβασης: *Νοέμβριος 2006 – Δεκέμβριος 2006*, (Κωδικός Προγράμματος 81472).
3. «Μετρική μελέτη και μελέτη επιδόσεων ελεύθερου λογισμικού ανοικτού κώδικα», *Έρευνα στην Τεχνολογία Λογισμικού*, Φορέας: *Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης Τμήμα Πληροφορικής*, Διάστημα σύμβασης: *Νοέμβριος 2006 – Δεκέμβριος 2006*, (Κωδικός Προγράμματος 81473).
4. «Αναμόρφωση του υπάρχοντος προγράμματος προπτυχιακών σπουδών του τμήματος Πληροφορικής», *Στατιστική Ανάλυση Δεδομένων*, Φορέας: *Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης Τμήμα Πληροφορικής*, Διάστημα σύμβασης: *Απρίλιος 2007 – Σεπτέμβριος 2007*, (Κωδικός Προγράμματος 21537).

5. «Υβριδικές Τεχνολογίες Διαχωρισμού», *Τεχνοοικονομική Ανάλυση*, Φορέας: *Τεχνολογικό Επαγγελματικό Ίδρυμα Καβάλας*, Διάστημα σύμβασης: *Μάιος 2007 - Αύγουστος 2007*, (Κωδικός Προγράμματος 06-E-003).
6. «Υβριδικές Τεχνολογίες Διαχωρισμού», *Τεχνοοικονομική Ανάλυση*, Φορέας: *Τεχνολογικό Επαγγελματικό Ίδρυμα Καβάλας*, Διάστημα σύμβασης: *Σεπτέμβριος 2007 - Μάρτιος 2008*, (Κωδικός Προγράμματος 06-E-003).
7. «Υπηρεσίες για τη σύνδεση προπτυχιακών και μεταπτυχιακών φοιτητών με την αγορά», *Τεχνική υποστήριξη*, Φορέας: *Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης Τμήμα Πληροφορικής*, Διάστημα σύμβασης: *Ιούνιος 2007 - Ιούλιος 2007*, (Κωδικός προγράμματος 21476).
8. «Στατιστική επεξεργασία δεδομένων για εντοπισμό εκπαιδευτικών αναγκών προνοιακού τομέα (ΣΤΕΠ)», *Εισαγωγή δεδομένων*, Φορέας: *Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης Τμήμα Πληροφορικής*, Διάστημα σύμβασης: *Σεπτέμβριος 2007 - Νοέμβριος 2007*, (Κωδικός προγράμματος 82781).
9. «Αναμόρφωση του υπάρχοντος Προγράμματος Προπτυχιακών Σπουδών του Τμήματος Πληροφορικής», *Στατιστική Ανάλυση Δεδομένων*, Φορέας: *Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης Τμήμα Πληροφορικής*, Διάστημα σύμβασης: *Οκτώβριος 2007 – Δεκέμβριος 2007*, (Κωδικός Προγράμματος 21537).
10. «Αναμόρφωση του υπάρχοντος Προγράμματος Προπτυχιακών Σπουδών του Τμήματος Πληροφορικής», *Στατιστική Ανάλυση Δεδομένων*, Φορέας: *Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης Τμήμα Πληροφορικής*, Διάστημα σύμβασης: *Ιανουάριος 2008 – Φεβρουάριος 2008*, (Κωδικός Προγράμματος 21537).
11. «Υβριδικές Τεχνολογίες Διαχωρισμού», *Τεχνοοικονομική Ανάλυση*, Φορέας: *Τεχνολογικό Επαγγελματικό Ίδρυμα Καβάλας*, Διάστημα σύμβασης: *Μάρτιος 2008 - Αύγουστος 2008*, (Κωδικός Προγράμματος 06-E-003).
12. «Υπηρεσίες για τη σύνδεση προπτυχιακών και μεταπτυχιακών φοιτητών με την αγορά», *Τεχνική υποστήριξη*, Φορέας: *Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης Τμήμα Πληροφορικής*, Διάστημα σύμβασης: *Απρίλιος 2008 - Ιούνιος 2008*, (Κωδικός προγράμματος 21476).
13. «Υπηρεσίες για τη σύνδεση προπτυχιακών και μεταπτυχιακών φοιτητών με την αγορά», *Τεχνική υποστήριξη*, Φορέας: *Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης Τμήμα Πληροφορικής*, Διάστημα σύμβασης: *Απρίλιος 2009 - Ιούνιος 2009*, (Κωδικός προγράμματος 21476).

ΕΡΕΥΝΗΤΙΚΟ ΕΡΓΟ

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΜΗΤΤΑΣ Ν. (2009). Στατιστικές και Υπολογιστικές Μέθοδοι Ανάπτυξης, Βελτίωσης και Σύγκρισης Μοντέλων Πρόβλεψης Κόστους Λογισμικού. *Διδακτορική Διατριβή του Τμήματος Πληροφορικής της Σχολής Θετικών Επιστημών, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης.*

ΔΗΜΟΣΙΕΥΣΕΙΣ ΣΕ ΔΙΕΘΝΗ ΠΕΡΙΟΔΙΚΑ

- J1.** **MITTAS Ν., ATHANASIADES Μ., ANGELIS L.** (2008). Improving Analogy – Based Software Cost Estimation by a Resampling Method. *Information and Software Technology (Elsevier), 50, 221–230.*
- J2.** **MITTAS Ν., ANGELIS L.** (2008). Comparing Cost Prediction Models by Resampling Techniques. *Journal of Systems and Software (Elsevier). Special Issue on Software Process and Product Measurement, 81, 616–632.*
- J3.** **MITTAS Ν., ANGELIS L.** (2010). Visual Comparison of Software Cost Estimation Models by Regression Error Characteristic Analysis. *Journal of Systems and Software (Elsevier), 83, 621-637.*
- J4.** **MITTAS Ν., ANGELIS L.** (2010). LSEbA: Least Squares Regression and Estimation by Analogy in a Semi-Parametric Model for Software Cost Estimation. *Empirical Software Engineering (Springer), 15 (5), 523-555.*
- J5.** **MITTAS Ν., ANGELIS L.** (2011). A Permutation Test based on Regression Error Characteristic Curves for Software Cost Estimation Models. *Empirical Software Engineering (Springer). Special Issue on Repeatable Results in Effort Estimation, 17 (1-2), 34-61.*
- J6.** **MITTAS Ν.** (2012). Evaluating the Performances of Software Cost Estimation Models through Prediction Intervals. *Journal of Engineering Science and Technology Review (accepted for publication).*

ΔΗΜΟΣΙΕΥΣΕΙΣ ΣΕ ΚΕΦΑΛΑΙΑ ΒΙΒΛΙΩΝ

- B1.** **ANGELIS L., SENTAS Π., MITTAS Ν., CHATZIPETROU Π.** (2010). Methods for Statistical and Visual Comparison of Imputation Methods for Missing Data in Software

Cost Estimation. In *Modern Software Engineering Concepts and Practices: Advanced Approaches*, Idea Group Inc. Publishing.

ΔΗΜΟΣΙΕΥΣΕΙΣ ΣΕ ΠΡΑΚΤΙΚΑ ΔΙΕΘΝΩΝ ΣΥΝΕΔΡΙΩΝ

- C1.** BIBI M., MITTAS N., ANGELIS L., STAMELOS I., MENDES E. (2007). Comparing Cross- vs. Within-Company Effort Estimation Models Using Interval Estimates. *IWSM-Mensura 2007 (International Conference on Software Process and Product Measurement)*, Palma de Mallorca, Spain, 5-8 November 2007.
- C2.** MITTAS N., ANGELIS L. (2008). Partial Regression Error Characteristic Curves for the Comparison of Software Cost Prediction Models, *Workshop on Artificial Intelligence, Techniques in Software Engineering (AISEW 2008)*, July 2008, Patras, Greece.
- C3.** MITTAS N., ANGELIS L. (2008). Comparing Software Cost Prediction Models by a Visualization Tool. *34th EUROMICRO Conference on Software Engineering and Advanced Applications (SEAA)*, September 2008, Parma, Italy.
- C4.** MITTAS N., ANGELIS L. (2008). Combining Regression and Estimation by Analogy in a Semi-parametric Model for Software Cost Estimation. *International Symposium on Empirical Software Engineering and Measurement ESEM'08*, October 9–10, 2008, Kaiserslautern, Germany.
- C5.** MITTAS N., ANGELIS L. (2009). Bootstrap Confidence Intervals for Regression Error Characteristic Curves Evaluating the Prediction Error of Software Cost Estimation Models. *2nd Artificial Intelligence Techniques in Software Engineering Workshop (AISEW2009)*, at the *5th IFIP Conference on Artificial Intelligence Applications and Innovations*, April, Thessaloniki, Greece. *Proceedings online*, <http://ceur-ws.org/Vol-475>, pp. 221-230.
- C6.** MITTAS N., ANGELIS L. (2009). Bootstrap Prediction Intervals for a Semi-Parametric Software Cost Estimation Model. *35th Euromicro Conference on Software Engineering and Advanced Applications (SEAA 2009)*. August 2009, Patras, Greece, *Proceedings published by IEEE*, pp. 293-299.
- C7.** MITTAS N., ARGYROPOYLOY V. ANGELIS L. (2010). Modeling the Relationship between Software Effort and Size Using Deming Regression. *6th International Conference on Predictive Models in Software Engineering*. September 12-13, 2010, Timisoara, Romania.
- C8.** KOSTI M., MITTAS N., ANGELIS L. (2010). DD-EbA: An Algorithm for Determining the Number of Neighbors in Cost Estimation by Analogy Using Distance Distributions. *Workshop on Artificial Intelligence, Techniques in Software Engineering (AISEW 2010)*, September 2010, Ayia Napa, Cyprus.

- C9.** HAMDAN K., ANGELIS L. **ΜΙΤΤΑΣ** N. (2010). Estimating learning by analogy: A case study in the UAE university. *International Conference of Education, Research and Innovation (ICERI 2010). November 15-17, 2010, Madrid, Spain.*
- C10.** **ΜΙΤΤΑΣ** N. (2011) Evaluating the Performances of Software Cost Estimation Models through Prediction Intervals. *International Conference on Econophysics, June 2-3, 2011, Kavala, Greece.*

ΔΗΜΟΣΙΕΥΣΕΙΣ ΣΕ ΠΡΑΚΤΙΚΑ ΕΘΝΙΚΩΝ ΣΥΝΕΔΡΙΩΝ - ΠΕΡΙΟΔΙΚΑ

- G1.** **ΜΗΤΤΑΣ** N., ΑΓΓΕΛΗΣ Ε. (2006). Διαστήματα Εμπιστοσύνης και Έλεγχοι Υποθέσεων στη Σύγκριση Μεθόδων Εκτίμησης Κόστους Λογισμικού. *Πρακτικά 19ου Πανελληνίου Συνεδρίου Στατιστικής, Καστοριά.*
- G2.** DIMOKAS N., **ΜΙΤΤΑΣ** N., NANOPOULOS A., ANGELIS L. (2008). A Prototype System for Educational Data Warehousing and Mining. *12th Pan-Hellenic Conference on Informatics PCI 2008 August 2008, Samos, Proceedings published by IEEE.*
- G3.** **ΜΗΤΤΑΣ** N., ΑΓΓΕΛΗΣ Ε. (2009). Γραφική Σύγκριση Μοντέλων Εκτίμησης Κόστους Λογισμικού Με Χαρακτηριστική Ανάλυση Σφάλματος Παλινδρόμησης. *Πρακτικά 22ου Πανελληνίου Συνεδρίου Στατιστικής, Χανιά.*
- G4.** **ΜΗΤΤΑΣ** N., ΦΛΩΡΟΥ Γ., ΠΟΛΥΧΡΟΝΙΔΟΥ Π. (2009). Διάρκεια Φοίτησης στο Τμήμα Λογιστικής ΤΕΙ Καβάλας μέσω της Ανάλυσης Επιβίωσης. *Πρακτικά 22ου Πανελληνίου Συνεδρίου Στατιστικής, Χανιά.*
- G5.** POLYCHRONIDOU P., **ΜΙΤΤΑΣ** N., FLOROU G. (2009). Factors that Affect the Duration of Studies of the Accounting Department of the Technological Institution of Kavala. *5th Pan-Hellenic Data Analysis Conference with International Participation, September 2009, Rethimno.*
- G6.** ΣΑΛΤΑΣ Β., ΚΑΛΑΜΠΑΚΑΣ Α., ΚΟΓΚΕΤΣΩΦ Α., ΠΟΛΥΧΡΟΝΙΔΟΥ Π., **ΜΗΤΤΑΣ** N., ΤΣΙΑΝΤΟΣ Β. (2010). Αξιολόγηση του Επιπέδου των Μαθηματικών των Πρωτοετών Φοιτητών της Σχολής Τεχνολογικών Εφαρμογών του ΤΕΙ Καβάλας. *Πρακτικά 23ου Πανελληνίου Συνεδρίου Στατιστικής, Βέροια.*
- G7.** ΤΣΙΑΝΤΟΣ Β., ΚΑΠΕΝΗΣ Κ., ΠΑΤΣΙΛΙΑΣ Γ., **ΜΗΤΤΑΣ** N., ΧΑΤΖΗΦΩΤΙΟΥ Σ. (2011). Η χρήση του Βίντεο στη Διδασκαλία των Μαθηματικών: Μία Πιλοτική Εφαρμογή στο Μάθημα «Μαθηματικά ΙΙ» του Τμήματος Βιομηχανικής Πληροφορικής του ΤΕΙ Καβάλας. *Πρακτικά 2^{ου} Πανελληνίου Συνεδρίου Ένταξη των ΤΠΕ στην Εκπαιδευτική Διαδικασία, Μάρτιος 2011, Πάτρα.*
- G8.** ΤΣΙΑΝΤΟΣ Β., **ΜΗΤΤΑΣ** N., ΠΟΛΥΧΡΟΝΙΔΟΥ Π., ΣΑΛΤΑΣ Β. (2011). Ανάλυση της Εξέλιξης των Φοιτητών του Τμήματος της Βιομηχανικής Πληροφορικής του ΤΕΙ Καβάλας

στα Μαθηματικά. Πρακτικά 1^{ου} Εθνικού Συνεδρίου της Ελληνικής Μαθηματικής Εταιρείας και της Ελληνικής Εταιρείας Επιχειρησιακών Ερευνών, Ιούνιος 2011, Πειραιάς.

ΠΡΟΣΩΠΙΚΗ ΣΥΜΜΕΤΟΧΗ ΣΕ ΣΥΝΕΔΡΙΑ

C2, C3, G1, G3, G4, C10.

ΠΛΗΡΟΦΟΡΙΑΚΑ ΣΤΟΙΧΕΙΑ ΔΗΜΟΣΙΕΥΣΕΩΝ

Impact Factor περιοδικών:

| Περιοδικό | IF (2010) | 5-Year IF |
|--|--------------|--------------|
| Journal of Systems and Software (Elsevier) | 1.227 | 1.290 |
| Information and Software Technology (Elsevier) | 1.507 | 1.426 |
| Empirical Software Engineering (Springer) | 1.776 | 1.783 |

Ποσοστό δεκτών εργασιών (acceptance rate) συνεδρίων

(ενδεικτικά για όσα συνέδρια βρέθηκαν τέτοια στοιχεία):

| Συνέδριο | Ποσοστό |
|--|---------|
| 2 nd ACM-IEEE international symposium on Empirical software engineering and measurement (ESEM 2008) | 28% |
| 35th Euromicro Conference on Software Engineering and Advanced Applications (SEAA 2009) | 39% |
| 12th Panhellenic Conference on Informatics 2008 | 49% |
| 34th Euromicro Conference on Software Engineering and Advanced Applications (SEAA 2008) | 56% |

ΚΡΙΤΗΣ ΣΕ ΠΕΡΙΟΔΙΚΑ

Κριτής στα περιοδικά:

- Journal of Systems and Software (Elsevier).
- IEEE Transactions on Software Engineering (IEEE).

ΚΡΙΤΗΣ ΣΕ ΔΙΕΘΝΗ ΣΥΝΕΔΡΙΑ

- 2nd IEEE International Conference on Computer and Communication Technology (ICCT - 2011), 15-17 September 2011 in Allahabad, India.

- Academic Program Committee on 12th EANN / 7th AIAI Joint Conferences, 15 - 18 September 2011, Corfu, Greece Engineering Applications of Neural Networks / Artificial Intelligence Applications and Innovations.
- Academic Program Committee on CISE 2012: 2nd International Workshop on Computational Intelligence in Software Engineering September 27-30, 2012 Halkidiki, Greece.

ΑΛΛΕΣ ΔΡΑΣΤΗΡΙΟΤΗΤΕΣ-ΕΠΙΣΤΗΜΟΝΙΚΗ ΕΠΙΜΟΡΦΩΣΗ

- Η Δυναμική της Σύγχρονης και Ασύγχρονης Τηλεκπαίδευσης στην Επαγγελματική Κατάρτιση. Σχέδια Μαθήματος. Πανεπιστήμιο Λισαβόνας, Πορτογαλία, Νοέμβριος 2006.
- Επιμόρφωση σε Θέματα Ισότητας των Φύλων. Κέντρο Ερευνών για Θέματα Ισότητας, 2006-2007.

ΔΙΕΘΝΕΙΣ ΣΥΝΕΡΓΑΣΙΕΣ

The University of Auckland (New Zealand)

- Emilia Mendes (Associate Professor) C1

ΕΤΕΡΟΑΝΑΦΟΡΕΣ ΣΤΟ ΕΡΕΥΝΗΤΙΚΟ ΕΡΓΟ

Σύνολο Ετεροαναφορών: 35

Αναφορές στην εργασία J1

- J1-1.** (2008) LI Q., WANG Q., YANG Y., LI M. Reducing biases in individual software effort estimations: a combining approach. *Proceedings of the Second ACM-IEEE international symposium on Empirical software engineering and measurement*, 223-232.
- J1-2.** (2008) KAMEI Y., KEUNG J., MONDEN A., MATSUMATO K. An over-sampling method for analogy-based software effort estimation. *Proceedings of the Second ACM-IEEE international symposium on Empirical software engineering and measurement*, 312-314
- J1-3.** (2008) LI Y.F., XIE M., GOH T.N. Optimization of feature weights and number of neighbors for analogy based cost estimation in software project management. *IEEE International Conference on Industrial Engineering and Engineering Management, IEEM 2008, art. no. 4738130, pp. 1542-1546*
- J1-4.** (2008) LI Y.F., XIE M., GOH T.N. A study of analogy based sampling for interval based cost estimation for software project management *Proceedings of the 4th IEEE International Conference on Management of Innovation and Technology, ICMIT, art. no. 4654377, pp. 281-286*

- J1-5.** (2008) LI Y.F., XIE M., GOH T.N. A bayesian inference approach for probabilistic analogy based software maintenance effort estimation *Proceedings of the 14th IEEE Pacific Rim International Symposium on Dependable Computing, PRDC 2008, art. no. 4725294, pp. 176-183*
- J1-6.** (2009) AZZEH M., NEAGU D., COWLING P. Software effort estimation based on weighted fuzzy grey relational analysis. *Proceedings (ACM) of the 5th International Conference on Predictor Models in Software Engineering table of contents Vancouver, British Columbia, Canada.*
- J1-7.** (2009) AZZEH M., NEAGU D., COWLING P. Fuzzy grey relational analysis for software effort estimation. *Empirical Software Engineering, 15(1), 60-90.*
- J1-8.** (2009) ZHANG J.G. Validity verifying method of software project management. *Proceedings of the International Conference on Management and Service Science, MASS 2009.*
- J1-9.** (2010) BO J., XIAOJUN Z., LIFENG X. Back propagation neural network based product cost estimation at an early design stage of passenger vehicles. *International Journal of Industrial and Systems Engineering, 5 (2), 190 – 211.*
- J1-10.** (2010) LI Y.F. Improvement and implementation of analogy based method for software project cost estimation. *Doctoral Thesis. Department of Industrial and Systems Engineering, Wuhan National University of Singapore.*
- J1-11.** (2011) WEN J., LI S., LIN Z, HU Z., HUANQ C. Systematic literature review of machine learning based software development effort estimation models. *Information and Software Technology.*
- J1-12.** (2012) NAGPAL G., UDDIN M, KAUR A. A Hybrid Technique using Grey Relational Analysis and Regression for Software Effort Estimation using Feature Selection. *International Journal of Software Computing and Engineering (IJSCE) 1 (6), January 2012, pp. 20-27.*

Αναφορές στην εργασία J2

- J2-1.** (2009) KITCHENHAM B., MENDES E. Why comparative effort prediction studies may be invalid. *Proceedings of the 5th International Conference on Predictor Models in Software Engineering table of contents, Vancouver, British Columbia, Canada.*
- J2-2.** (2009) VIJAY J.F., MANOHARAN C. Initial Hybrid Method for Analyzing Software Estimation, Benchmarking and Risk Assessment Using Design of Software. *Journal of Computer Science 5 (10): 717-724.*
- J2-3.** (2010) ARSHID A., SALMAN Q., SYED SHAH M., JALIL A., MUHAMMAD TARIQPERVA., SARFARAZ A. Software Cost Estimation through Entity Relationship Model. *Journal of American Science, 6 (11), 47-51.*
- J2-4.** (2010) KHAN K. The evaluation of well-known effort estimation models based on predictive accuracy indicators. *Master Thesis. Number: MSE- 2010:03, School of Computing Blekinge Institute of Technology, Sweden.*

- J2-5.** (2010) PAHARIYAA J., RAVIA, V., CARRA M., VASUA M. Computational Intelligence Hybrids Applied to Software Cost Estimation. *International Journal of Computer Information Systems and Industrial Management Applications 2*, 104-112.
- J2-6.** (2012) FALESSI D., CANTONE G., CANFORA G. Empirical Principles and an Industrial Case Study in Retrieving Equivalent Requirements via Natural Language Processing Techniques. *IEEE Transactions on Software Engineering*.
- J2-7.** (2012) NAGPAL G., UDDIN M, KAUR A. A Hybrid Technique using Grey Relational Analysis and Regression for Software Effort Estimation using Feature Selection. *International Journal of Software Computing and Engineering (IJSCE) 1 (6), January 2012, pp. 20-27.*

Αναφορές στην εργασία J3

- J3-1.** (2010) FERNADEZ-DIEGO M., MARTINEZ-GOMEZ M. TORRABLA-MARTINEZ J.M. Sensitivity of results to different data quality meta-data criteria in the sample selection of projects from the ISBSG dataset. *Proceedings of the 6th International Conference on Predictive Models in Software Engineering. September 12-13, 2010, Timisoara, Romania.*
- J3-2.** (2010) PAHARIYAA J., RAVIA, V., CARRA M., VASUA M. Computational Intelligence Hybrids Applied to Software Cost Estimation. *International Journal of Computer Information Systems and Industrial Management Applications 2*, 104-112.

Αναφορές στην εργασία J4

- J4-1.** (2011) S. GUPTA, G. SIKKA, H. VERMA. Recent methods for software effort estimation by analogy. *ACM SIGSOFT Software Engineering Notes*, 36 (4), 1-5.
- J4-2.** (2011) WEN J., LI S., LIN Z, HU Z., HUANQ C. Systematic literature review of machine learning based software development effort estimation models. *Information and Software Technology*.
- J4-3.** (2012) NAGPAL G., UDDIN M, KAUR A. A Hybrid Technique using Grey Relational Analysis and Regression for Software Effort Estimation using Feature Selection. *International Journal of Software Computing and Engineering (IJSCE) 1 (6), January 2012, pp. 20-27.*

Αναφορές στην εργασία J5

- J5-1.** (2012) MENZIES T., SHEPPERD M. Special issue on repeatable results in software engineering prediction. *Empirical Software Engineering (Springer). Special Issue on Repeatable Results in Effort Estimation.*

Αναφορές στην εργασία C4

- C4-1.** (2009) SEO Y.-S., YOON K.-A., BAE, D.-H. Improving the accuracy of software effort estimation based on multiple least square regression models by estimation error-based data partitioning. *Proceedings - Asia-Pacific Software Engineering Conference, APSEC*: pp. 3-10.
- C4-2.** (2009) ASSUNÇÃO S.C. Caracterização da prática da gestão de projectos de desenvolvimento de software - Perspectiva dos especialistas de prestadores de serviços. DISSERTAÇÃO DE MESTRADO EM TECNOLOGIAS DA INFORMAÇÃO E COMUNICAÇÃO. UNIVERSIDADE DE TRÁS-OS-MONTES E ALTO DOURO.
- C4-3.** (2010) KHAN K. The evaluation of well-known effort estimation models based on predictive accuracy indicators. *Master Thesis. Number: MSE-2010:03, School of Computing Blekinge Institute of Technology, Sweden.*

Αναφορές στην εργασία C6

- C6-1.** (2010) KHAN K. The evaluation of well-known effort estimation models based on predictive accuracy indicators. *Master Thesis. Number: MSE-2010:03, School of Computing Blekinge Institute of Technology, Sweden.*
- C6-2.** (2011) KLAS M., TRENDOWICZ A., ISHIGAI Y., NAKAO H. Handling Estimation Uncertainty with Bootstrapping: Empirical Evaluation in the Context of Hybrid Prediction Methods. *Proceedings of the 5th International Symposium on Empirical Software Engineering and Measurement (ESEM 2011), Sept. 22-23, Banff, Canada.*

Αναφορές στην εργασία C7

- C7-1.** (2010) PAPTHEOCHAROUS E., ANDREOU A. Size-based software cost modelling with artificial neural networks and genetic algorithms. *Artificial Neural Networks – Application, InTech*, pp. 167-188, April 2011.

Αναφορές στην εργασία C8

- C8-1.** (2011) S. GUPTA, G. SIKKA, H. VERMA. Recent methods for software effort estimation by analogy. *ACM SIGSOFT Software Engineering Notes*, 36 (4), 1-5.

Αναφορές στην εργασία G2

- G2-1.** (2011) AREVALILLO-HERAEZ M., MORENO-CLARI P., CERVERON-LLEO V. Educational knowledge generation from administrative data. *Journal of Educational Technology Research and Development (Springer)*.

- G2-2.** (2010) WARNARS S. Tata Kelola Database Perguruan Tinggi Yang Optimal Dengan Data Warehouse. *TELKOMNIKA Indonesian Journal of Electrical Engineering*, 8 (1), pp. 25 – 34.
- G2-3.** (2011) Y. Xuejian, L. Xueqing. A Multidimensional Data Analysis System Based on MDA for Educational Data Warehousing. *Proceedings of the 6th International Conference on Computer Science & Education (ICCSE 2011)*, August 3-5, SuperStar Virgo, Singapore, pp. 88-94.

ΝΙΚΟΛΑΟΣ Α. ΜΗΤΤΑΣ

ΥΠΟΜΝΗΜΑ ΕΡΓΑΣΙΩΝ

ΘΕΣΣΑΛΟΝΙΚΗ 2011

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΜΗΤΤΑΣ Ν. (2009). Στατιστικές και Υπολογιστικές Μέθοδοι Ανάπτυξης, Βελτίωσης και Σύγκρισης Μοντέλων Πρόβλεψης Κόστους Λογισμικού. Διδακτορική Διατριβή του Τμήματος Πληροφορικής της Σχολής Θετικών Επιστημών, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης.

Η *Εκτίμηση Κόστους Λογισμικού* (Software Cost Estimation-SCE) αποτελεί μία από τις σπουδαιότερες δραστηριότητες για την παράδοση έργων λογισμικού εντός των προκαθορισμένων χρονοδιαγραμμάτων και προϋπολογισμών με την απαιτούμενη υψηλή ποιότητα. Στα πρώτα στάδια του κύκλου ζωής, η εκτίμηση του κόστους σχετίζεται άμεσα με τις δραστηριότητες της σχεδίασης ενός έργου ενώ στα επόμενα στάδια, υποστηρίζει τις διοικητικές δραστηριότητες ενός οργανισμού για το σχεδιασμό, χρονοπρογραμματισμό και την ανάλυση κινδύνου. Λόγω της αυξανόμενης ζήτησης για έργα λογισμικού με υψηλές προδιαγραφές, οι οργανισμοί αναζητούν τεχνικές εκτίμησης του κόστους με τις οποίες θα διασφαλίζουν ότι τα έργα που παράγουν ακολουθούν τις επιτακτικές ανάγκες της σημερινής εποχής.

Η ερευνητική προσπάθεια στα πλαίσια της παρούσας διατριβής επικεντρώνεται στην εισαγωγή και μελέτη στατιστικών και υπολογιστικών μεθόδων με σκοπό τη στατιστική σύγκριση των εναλλακτικών μεθόδων εκτίμησης, τη βελτίωση μίας γνωστής τεχνικής εκτίμησης και τέλος, την ανάπτυξη μίας νέας μεθοδολογίας μοντελοποίησης του κόστους που καταφέρνει να συνδυάσει δύο από τα πιο γνωστά μοντέλα του συγκεκριμένου ερευνητικού πεδίου.

Αρχικά, γίνεται προσπάθεια να αντιμετωπιστεί το διοικητικό πρόβλημα των οργανισμών, οι οποίοι καλούνται να επιλέξουν ανάμεσα από ένα πλήθος τεχνικών εκτίμησης, το βέλτιστο μοντέλο που θα χρησιμοποιηθεί για τη δύσκολη δραστηριότητα της πρόβλεψης ενός νέου έργου λογισμικού. Λόγω των προβλημάτων των υπαρχουσών στρατηγικών επιλογής, στην παρούσα διατριβή προτείνονται ανθεκτικές στατιστικές τεχνικές προσομοίωσης, όπως η κατασκευή διαστημάτων εμπιστοσύνης με τη χρήση της *bootstrap αναδειγματοληψίας* (resampling) και οι *έλεγχοι μεταθέσεων* (permutation tests). Οι παραπάνω τεχνικές καλούνται να αντιμετωπίσουν προβλήματα των κλασικών προσεγγίσεων, όπως ο μικρός αριθμός των έργων, η μη-κανονικότητα και η μεγάλη λοξότητα των σφαλμάτων πρόβλεψης.

Παράλληλα, προτείνεται ένα γραφικό πλαίσιο εργασίας, η ανάλυση με *Χαρακτηριστικές Καμπύλες Σφάλματος Παλινδρόμησης* (Regression Error Characteristic-REC Curves) με ενδιαφέρουσες γεωμετρικές ιδιότητες, η οποία έχει στόχο τη βελτίωση της διαδικασίας *αξιολόγησης* (validation) των μοντέλων και της *σύγκρισης* (comparison) εναλλακτικών μεθόδων εκτίμησης μέσω της απλής γραφικής απεικόνισης των σφαλμάτων πρόβλεψης. Η αξιοποίηση των πολλαπλών δυνατοτήτων του προτεινόμενου γραφικού εργαλείου είναι ικανή να διευκολύνει τους διοικητές έργου να αποκτήσουν σημαντική πληροφορία τόσο για την αξιολόγηση μίας τεχνικής εκτίμησης, όσο και για τη διαδικασία σύγκρισης εναλλακτικών μοντέλων. Βάσει των παραπάνω, η προτεινόμενη μεθοδολογία συντελεί στην ορθή καθοδήγηση των διεργασιών κατά τη διάρκεια της ανάπτυξης των έργων και την τελική επιλογή του αποτελεσματικότερου μοντέλου πρόβλεψης του κόστους.

Το κομμάτι που αναφέρεται στη σύγκριση των μοντέλων εκτίμησης ολοκληρώνεται με την εξέταση ενός σημαντικού ερευνητικού ζητήματος που αφορά τη σύγκριση μοντέλων που βασίζονται σε δεδομένα ενός ή πολλών οργανισμών για την πρόβλεψη του κόστους ενός νέου έργου. Το παραπάνω πρακτικό πρόβλημα έχει απασχολήσει ιδιαίτερα, την ερευνητική κοινότητα αλλά οι ως τώρα προσπάθειες έχουν επικεντρωθεί στη σύγκριση τεχνικών που βασίζονται σε *σημειακές εκτιμήσεις* (point estimates). Η παρούσα διατριβή επεκτείνει την υπάρχουσα κατάσταση σε συγκρίσεις μεθόδων που βασίζονται σε *εκτιμήσεις διαστήματος* (interval estimates) λαμβάνοντας υπόψη την αβεβαιότητα και το ρίσκο. Στην προσπάθεια αυτήν, συγκρίνονται οι προβλέψεις που προέρχονται από δύο διαφορετικούς τύπους τεχνικών: (α) τεχνικές που παράγουν εκτιμήσεις σε σημείο και συνοδεύονται από ένα διάστημα και (β) τεχνικές που παράγουν προκαθορισμένα διαστήματα, από τα οποία υπολογίζεται η σημειακή εκτίμηση. Τέλος, αναπτύσσεται μία νέα μεθοδολογία για τη σύγκριση των εκτιμήσεων σε διάστημα, η

οποία λαμβάνει υπόψη την ικανότητα μίας τεχνικής να προβλέπει στο σωστό διάστημα το κόστος ενός έργου αλλά παράλληλα ενσωματώνει σημαντική πληροφορία που σχετίζεται με το εύρος και την αλληλοεπικάλυψη των υπό σύγκριση διαστημάτων.

Ακολούθως, εξετάζεται το πρόβλημα της βελτίωσης της μεθόδου *Εκτίμησης με Αναλογίες* (Estimation by Analogy-EbA), η οποία έχει προσελκύσει το ενδιαφέρον πολλών ερευνητών του χώρου αλλά από την άλλη μεριά, παρουσιάζει αρκετά μειονεκτήματα λόγω της εμπειρικής της φύσης. Στην παρούσα διατριβή, αρχικά γίνεται προσπάθεια για τη μοντελοποίηση ολόκληρης της διαδικασίας, αξιοποιώντας την άμεση σχέση της τεχνικής με τη *μη-παραμετρική παλινδρόμηση* (non-parametric regression). Παράλληλα, βάσει της παραπάνω σχέσης προτείνεται η μεθοδολογία αναδειγματοληψίας *iterated bagging* με στόχο τη μείωση του σφάλματος πρόβλεψης. Το σημαντικότερο πλεονέκτημα της μεθόδου είναι ότι καταφέρνει να αντιμετωπίσει το πρόβλημα του μικρού αριθμού έργων ενός συνόλου δεδομένων, ενώ είναι ικανή να προβλέψει το κόστος εκτός του εύρους τιμών των παλαιότερων έργων της βάσης.

Τέλος, αναπτύσσεται μία νέα *ημι-παραμετρική* (semi-parametric) στατιστική μεθοδολογία (*Least Squares Estimation by Analogy-LSEbA*) για τη μοντελοποίηση του κόστους των έργων, συνδυάζοντας τα πλεονεκτήματα της παραμετρικής Ανάλυσης Παλινδρόμησης και της μη-παραμετρικής Εκτίμησης με Αναλογίες. Η προτεινόμενη τεχνική φαίνεται να αποτελεί μία ρεαλιστική προσέγγιση για τη διαδικασία της εκτίμησης λαμβάνοντας υπόψη τους διαφορετικούς τύπους των ανεξάρτητων μεταβλητών ενώ καταφέρνει να διαχειριστεί πολλά προβλήματα, όπως την ύπαρξη πολλών κατηγορικών μεταβλητών, τις ελλείπουσες τιμές και την ακρίβεια των διαστημάτων πρόβλεψης που συνοδεύουν την σημειακή εκτίμηση.

Συνοψίζοντας, η διατριβή αυτή προτείνει μια σειρά από μεθόδους οι οποίες μπορούν να γενικευτούν περαιτέρω και να εφαρμοστούν είτε στην πράξη για την ορθή διοίκηση των έργων λογισμικού, είτε στην έρευνα που στοχεύει στη βελτίωση της διαδικασίας εκτίμησης του κόστους.

ΔΗΜΟΣΙΕΥΣΕΙΣ ΣΕ ΔΙΕΘΝΗ ΠΕΡΙΟΔΙΚΑ

[J1] MITTAS N, ATHANASIADES M., ANGELIS L. (2008). Improving Analogy – Based Software Cost Estimation by a Resampling Method. *Information and Software Technology (Elsevier)*, 50, 221–230.

Η εκτίμηση με αναλογίες (Estimation by Analogy - EbA) είναι μια πολύ γνωστή και συνήθης τεχνική στην περιοχή εκτίμησης κόστους λογισμικού (Software Cost Estimation-SCE). Η δημοτικότητα της μεθόδου οφείλεται στην αμεσότητά της και στην διαισθητικά ελκυστική ερμηνεία της. Ωστόσο, παρά την απλότητά της στην εφαρμογή, η θεωρητική μελέτη της EbA είναι αρκετά περίπλοκη. Στην εργασία αυτή εκμεταλλευόμαστε τη στενή σχέση της EbA με τη στατιστική μεθοδολογία που είναι γνωστή ως μη-παραμετρική παλινδρόμηση πλησιέστερων γειτόνων (nearest neighbor non-parametric regression) με σκοπό να προτείνουμε μια μέθοδο επαναλαμβανόμενης δειγματοληψίας για τη μείωση του σφάλματος πρόβλεψης. Η βελτιωτική επίδραση της μεθόδου στην EbA αξιολογείται χρησιμοποιώντας τεχνητά και πραγματικά δεδομένα από τη βιβλιογραφία.

[J2] MITTAS N., ANGELIS L. (2008). Comparing Cost Prediction Models by Resampling Techniques. *Journal of Systems and Software (Elsevier)*, Special Issue on Software Process and Product Measurement, 81, 616–632.

Η ακρίβεια της εκτίμησης του κόστους έργων λογισμικού είναι ένα σημαντικό ερευνητικό θέμα στη διοίκηση έργων (project management) εδώ και αρκετές δεκαετίες. Το ερευνητικό ενδιαφέρον έχει επικεντρωθεί στην ανάπτυξη διαφόρων στατιστικών μοντέλων, τα οποία βασίζονται σε ιστορικά δεδομένα παλαιότερων έργων. Η πληθώρα των εναλλακτικών προσεγγίσεων, οι οποίες μπορούν να

προσαρμοστούν στα διάφορα σύνολα δεδομένων για την εκτίμηση κόστους, έχει αναδείξει τη σπουδαιότητα και την ιδιαίτερη σημασία της επιλογής του πιο αποδοτικού μοντέλου πρόβλεψης. Συνήθως, η πάγια στρατηγική που ακολουθείται στην πράξη από τους διοικητές των έργων αλλά και από τους ερευνητές, για την επιλογή του καταλληλότερου μοντέλου, βασίζεται στη σύγκριση συγκεκριμένων μέτρων ακρίβειας, τα οποία είναι συνήθως συναρτήσεις των σχετικών σφαλμάτων των εναλλακτικών μεθόδων πρόβλεψης.

Η προαναφερόμενη πρακτική στηρίζεται στην απλή αντιπαράθεση των μέτρων ακρίβειας χωρίς να διεξάγεται κανένας έλεγχος υποθέσεως για τη στατιστική σημαντικότητα των αποτελεσμάτων. Η συγκεκριμένη πολιτική μπορεί να οδηγήσει σε εσφαλμένα και ασταθή συμπεράσματα, από τη στιγμή που μία μικρή αλλαγή στα δεδομένα είναι ικανή να αντιστρέψει και να ανατρέψει το αρχικό συμπέρασμα. Επιπρόσθετα, οι δείκτες ακρίβειας στους οποίους βασίζεται η σύγκριση, αποτελούν στην ουσία στατιστικά μέτρα, δηλαδή τυχαίες μεταβλητές, που ακολουθούν άγνωστες κατανομές. Το γεγονός αυτό δυσχεραίνει τους ελέγχους υποθέσεων που βασίζονται σε κλασικές παραμετρικές προσεγγίσεις.

Στην εργασία αυτή, προτείνεται η χρήση στατιστικών μεθόδων προσομοίωσης, οι οποίες μπορούν να εφαρμοστούν στη διαδικασία σύγκρισης της ακρίβειας εναλλακτικών μοντέλων πρόβλεψης κόστους. Στην προσπάθεια αυτή, χρησιμοποιήθηκαν οι δύο πιο δημοφιλείς τεχνικές εκτίμησης κόστους, η Εκτίμηση με Αναλογίες και η Παλινδρόμηση Ελαχίστων Τετραγώνων αλλά τα αποτελέσματα της μεθοδολογίας μπορούν να εφαρμοστούν για συγκρίσεις οποιονδήποτε μεθόδων. Οι προτεινόμενες στατιστικές μέθοδοι προσομοίωσης βασίζονται στους ελέγχους μεταθέσεων (permutation tests) και την κατασκευή διαστημάτων εμπιστοσύνης για τη διαφορά των μέτρων ακρίβειας μέσω της bootstrap αναδειγματοληψίας. Στο πειραματικό κομμάτι της εργασίας, χρησιμοποιήθηκαν τρία γνωστά από την βιβλιογραφία σύνολα δεδομένων για να επικυρώσουμε τα αποτελέσματα και να δείξουμε τις διαφορές μεταξύ των κλασικών παραμετρικών και μη-παραμετρικών προσεγγίσεων με τις προτεινόμενες τεχνικές προσομοίωσης.

[J3] MITTAS N., ANGELIS L. (2010). Visual Comparison of Software Cost Estimation Models by Regression Error Characteristic Analysis. *Journal of Systems and Software (Elsevier)*, 83, 621-637.

Η ισορροπημένη διοίκηση ενός έργου λογισμικού βασίζεται σε μεγάλο βαθμό στην επιλογή του καταλληλότερου μοντέλου πρόβλεψης μέσα από ένα πλήθος μεθόδων που έχουν προταθεί στη βιβλιογραφία. Η επιλογή αυτή συνήθως βασίζεται σε μέτρα του σφάλματος των εκτιμήσεων ενώ σε πιο πρόσφατες μελέτες, οι ερευνητές κάνουν χρήση στατιστικών διαδικασιών. Η πρώτη προσέγγιση μπορεί να οδηγήσει σε μη σταθερά και εσφαλμένα συμπεράσματα εξαιτίας της ύπαρξης απόμακρων σημείων. Όσον αφορά τη δεύτερη προσέγγιση, οι στατιστικοί έλεγχοι θα πρέπει να εφαρμοστούν από έναν ειδικό με ισχυρό στατιστικό υπόβαθρο, ενώ παράλληλα τα αποτελέσματα δε μπορούν να παρουσιαστούν, να ερμηνευθούν και να γίνουν εύκολα αντιληπτά από μη-ειδικούς.

Στην εργασία αυτή προτείνεται η ανάλυση με Χαρακτηριστικές Καμπύλες Σφάλματος Παλινδρόμησης (Regression Error Characteristic Curves), ενός ισχυρού εργαλείου γραφικής απεικόνισης, με ενδιαφέρουσες γεωμετρικές ιδιότητες, για τη σύγκριση και αξιολόγηση των εναλλακτικών μοντέλων πρόβλεψης μέσω της απλής εξέτασης ενός γραφήματος. Επιπρόσθετα, προτείνεται ένα πλαίσιο ανάλυσης, το οποίο καλύπτει διάφορες πτυχές της διεργασίας εκτίμησης του κόστους, όπως τη βαθμονόμηση μίας μεθοδολογίας εκτίμησης, την αναγνώριση παραγόντων που επηρεάζουν το σφάλμα, τη συστηματική διερεύνηση σφαλμάτων σε προκαθορισμένα πεδία τιμών του πραγματικού κόστους και την εξέταση της κατανομής του κόστους σε συγκεκριμένα πεδία τιμών του σφάλματος. Εκτενή πειράματα σε ένα σύνολο με πραγματικά δεδομένα επιδεικνύουν με εμφανή τρόπο τα οφέλη, τα πλεονεκτήματα και τη σημαντική απόκτηση πληροφορίας, η οποία προκύπτει από την εισαγωγή της προτεινόμενης ανάλυσης.

[J4] MITTAS N., ANGELIS L. (2010). LSEbA: Least Squares Regression and Estimation by Analogy in a Semi-Parametric Model for Software Cost Estimation. *Empirical Software Engineering (Springer)*, 15 (5), 523-555.

Η σημαντικότητα της εκτίμησης κόστους λογισμικού στα πρώτα στάδια του κύκλου ζωής της ανάπτυξης απεικονίζεται ξεκάθαρα από τη χρήση πολλών μοντέλων και μεθόδων, οι οποίες έχουν εμφανιστεί μέχρι τώρα στη βιβλιογραφία. Το ενδιαφέρον των ερευνητών έχει επικεντρωθεί σε δύο πολύ γνωστές τεχνικές, οι οποίες είναι η παραμετρική ανάλυση παλινδρόμησης (parametric Regression Analysis) και η μη-παραμετρική εκτίμηση με αναλογίες (non-parametric Estimation by Analogy). Παρά το μεγάλο πλήθος των συγκριτικών μελετών, φαίνεται πως υπάρχει μία ασυμφωνία στην επιλογή της καλύτερης τεχνικής πρόβλεψης. Σε αυτή την εργασία, εισάγεται μία ημι-παραμετρική (semi-parametric) τεχνική, η οποία κατορθώνει να συνδυάσει τις δύο προαναφερόμενες μεθόδους, διατηρώντας τα πλεονεκτήματα κάθε τεχνικής. Επιπρόσθετα, η προτεινόμενη μεθοδολογία διαχειρίζεται κατάλληλα τους διάφορους τύπους δεδομένων και εκμεταλλεύεται ολόκληρη την αρχική πληροφορία του συνόλου των δεδομένων, ακόμα και σε περιπτώσεις ύπαρξης μεγάλου όγκου ελλειπουσών τιμών (missing values). Στην εργασία επεξηγείται αναλυτικά η διαδικασία δημιουργίας του προτεινόμενου μοντέλου και παρουσιάζονται εκτενή πειράματα σε διαφορετικά χαρακτηριστικά σύνολα δεδομένων, επιβεβαιώνοντας τα οφέλη της προτεινόμενης τεχνικής. Οι συγκρίσεις του προτεινόμενου μοντέλου με τη γραμμική παλινδρόμηση και την εκτίμηση με αναλογίες διεξάγονται μέσω των μέτρων ακρίβειας, των στατιστικών ελέγχων υποθέσεων και του γραφικού εργαλείου της ανάλυσης χαρακτηριστικών καμπυλών σφάλματος παλινδρόμησης.

[J5] MITTAS N., ANGELIS L. (2011). A Permutation Test based on Regression Error Characteristic Curves for Software Cost Estimation Models. *Empirical Software Engineering (Springer). Special Issue on Repeatable Results in Effort Estimation*, 17 (1-2), 34-61.

Οι Χαρακτηριστικές Καμπύλες Σφάλματος Παλινδρόμησης (Regression Error Characteristic Curves) αποτελούν ένα εργαλείο οπτικής απεικόνισης, το οποίο είναι κατάλληλο για να αναδείξει την ικανότητα πρόβλεψης εναλλακτικών μοντέλων. Οι παραπάνω καμπύλες πρόσφατα προτάθηκαν στην επιστημονική περιοχή της εκτίμησης κόστους λογισμικού (Software Cost Estimation-SCE) και αποτελούν ένα ισχυρό εργαλείο διοίκησης των έργων εξαιτίας των πλεονεκτημάτων που προέρχονται από τη χρήση τους και κυρίως λόγω της απεικόνισης ολόκληρης της κατανομής του σφάλματος που υπολογίζεται από τα εναλλακτικά μοντέλα πρόβλεψης. Παρά το γεγονός ότι παρέχουν στους διοικητές των έργων σημαντική πληροφορία μέσω ενός γραφήματος, οι REC καμπύλες δε μπορούν να απαντήσουν στο κρίσιμο ερώτημα για το πότε οι αποκλίσεις μεταξύ των διαφορετικών κατανομών σφάλματος μπορούν να αποτελέσουν ισχυρή ένδειξη για στατιστικά σημαντική διαφορά στην ακρίβεια πρόβλεψης των εναλλακτικών μοντέλων. Στην εργασία αυτή προτείνουμε μία γραφική διαδικασία που χρησιμοποιεί (α) επαναλαμβανόμενες αντιμεταθέσεις και (β) τη μέγιστη κάθετη απόκλιση μεταξύ δύο Χαρακτηριστικών Καμπυλών Σφάλματος Παλινδρόμησης ώστε να διεξάγουμε έναν έλεγχο υποθέσεως για να μπορέσουμε να συμπεράνουμε τη στατιστική σημαντικότητα της διαφοράς μεταξύ δύο συναρτήσεων σφάλματος. Τα αποτελέσματα καταδεικνύουν την αναγκαιότητα του προτεινόμενου ελέγχου υποθέσεως για να εξάγουμε σημαντικά συμπεράσματα για την ικανότητα πρόβλεψης εναλλακτικών μοντέλων κόστους λογισμικού. Επιπρόσθετα, η προτεινόμενη στατιστική διαδικασία μπορεί να εφαρμοστεί σε οποιαδήποτε σύνολα δεδομένων όπου στόχος είναι η πρόβλεψη μίας εξαρτημένης μεταβλητής και η σύγκριση διαφόρων μοντέλων πρόβλεψης ώστε να αναδειχθεί η καλύτερη στρατηγική που πρέπει να ακολουθήσει ο ερευνητής. Ο προτεινόμενος έλεγχος υποθέσεως αποτελεί στην πράξη ένα γραφικό εργαλείο, το οποίο μπορεί να ερμηνευθεί ευκολότερα από τις κλασικές παραμετρικές και μη-παραμετρικές στατιστικές διαδικασίες ενώ ακόμα είναι απαλλαγμένος από τις υποθέσεις της κανονικής κατανομής και δεν επηρεάζεται από το μικρό πλήθος των παρατηρήσεων και τη μεγάλη λοξότητα που παρουσιάζουν οι κατανομές των σφαλμάτων.

[J6] MITTAS N. (2012). Evaluating the Performances of Software Cost Estimation Models through Prediction Intervals. *Journal of Engineering Science and Technology Review* (accepted for publication).

Στην εργασία αυτή προτείνεται ένα στατιστικό πλαίσιο εργασίας, το οποίο επικεντρώνεται στην κατασκευή διαστημάτων πρόβλεψης. Ένα διάστημα πρόβλεψης παρέχει στον διοικητή έργου μία οπτιμιστική και πεσιμιστική εκτίμηση για την πραγματική τιμή του κόστους ενός έργου λογισμικού. Το προτεινόμενο πλαίσιο εργασίας, το οποίο ενσωματώνει διάφορους δείκτες ακρίβειας, τη διεξαγωγή ενός ελέγχου υποθέσεως και γραφική σύγκριση της απόδοσης πρόβλεψης εφαρμόζεται σε ένα σύνολο δεδομένων με πραγματικά έργα λογισμικού.

ΔΗΜΟΣΙΕΥΣΕΙΣ ΣΕ ΚΕΦΑΛΑΙΑ ΒΙΒΛΙΩΝ

[B1] ANGELIS L., SENTAS P., MITTAS N., CHATZIPETROU P (2010). Methods for Statistical and Visual Comparison of Imputation Methods for Missing Data in Software Cost Estimation. *In Modern Software Engineering Concepts and Practices: Advanced Approaches, Idea Group Inc. Publishing.*

Ένα σημαντικό πρόβλημα στην κατασκευή αξιόπιστων μοντέλων κόστους λογισμικού είναι το γεγονός ότι τα διαθέσιμα σύνολα περιέχουν πολλές ελλείπουσες παρατηρήσεις για κατηγορικά δεδομένα. Η εργασία αυτή έχει στόχο να αναδείξει τον τρόπο με τον οποίο μοντέρνες στατιστικές και υπολογιστικές μέθοδοι μπορούν να χρησιμοποιηθούν για τη σύγκριση της επίδρασης των τεχνικών συμπλήρωσης ελλειπουσών παρατηρήσεων στην ακρίβεια ενός μοντέλου του κόστους. Πιο συγκεκριμένα, γίνεται χρήση της Πολωνυμικής Λογιστικής Παλινδρόμησης (Multinomial Logistic Regression), η οποία συγκρίνεται με τέσσερις παλαιότερες τεχνικές. Η αξιολόγηση βασίζεται σε πολλές συναρτήσεις του σφάλματος πρόβλεψης και συγκρίσεις με τη βοήθεια στατιστικών ελέγχων υποθέσεων, τεχνικών αναδειγματοληψίας και των REC καμπύλων.

ΔΗΜΟΣΙΕΥΣΕΙΣ ΣΕ ΠΡΑΚΤΙΚΑ ΔΙΕΘΝΩΝ ΣΥΝΕΔΡΙΩΝ

[C1] BIBI M., MITTAS N., ANGELIS L., STAMELOS I., MENDES E. (2007). Comparing Cross-vs. Within-Company Effort Estimation Models Using Interval Estimates. *IWSM-Mensura 2007 (International Conference on Software Process and Product Measurement), Palma de Mallorca, Spain, 5-8 November 2007.*

Στην εργασία αυτή συγκρίνονται διαφορετικές τεχνικές εκτίμησης του κόστους. Το έναυσμα για τη συγκεκριμένη μελέτη αποτέλεσε το ερώτημα σχετικά με το κατά πόσο υπάρχουν διαφορές στην ακρίβεια πρόβλεψης διαφορετικών μοντέλων κόστους, όταν αυτά βασίζονται σε σύνολα δεδομένων από έναν οργανισμό ή σε δεδομένα από πολλούς οργανισμούς. Επιπρόσθετα, το ερώτημα επεκτείνεται ώστε να συμπεριλάβουμε στη διαδικασία της σύγκρισης και μοντέλα που βασίζονται τόσο σε ενδοεταιρικά αλλά και σε πολυσυλλεκτικά δεδομένα (μεικτά μοντέλα). Ορισμένες από τις τεχνικές που εξετάζονται παράγουν σημειακές εκτιμήσεις (point estimates), συνοδευμένες από ένα διάστημα πρόβλεψης (prediction interval) ενώ άλλες τεχνικές παράγουν προβλέψεις σε διάστημα (interval predictions), από όπου μπορεί να εκτιμηθεί μια πρόβλεψη σε σημείο. Κάθε μία από τις τεχνικές που εξετάζονται εφαρμόζεται στα ενδοεταιρικά, στα πολυσυλλεκτικά και στα μεικτά δεδομένα ώστε να ελεγχθεί με στατιστικό τρόπο, το σύνολο δεδομένων που αποτελεί τη βάση για ακριβέστερες εκτιμήσεις. Για τη διαδικασία της σύγκρισης των σημειακών εκτιμήσεων γίνεται χρήση αρκετών μέτρων του σφάλματος πρόβλεψης ενώ για τη σύγκριση των διαστημάτων πρόβλεψης προτείνεται και χρησιμοποιείται ένα νέο μέτρο, το οποίο αποτελεί γενίκευση του γνωστού μέτρου του ποσοστού επιτυχίας και λαμβάνει υπόψη (α) το κατά πόσο η πραγματική τιμή κόστους πέφτει μέσα σε ένα διάστημα, (β) την ομοιότητα των υπό

σύγκριση διαστημάτων και (γ) το εύρος των διαστημάτων. Η εκτίμηση με αναλογίες και η ανάλυση παλινδρόμησης χρησιμοποιούνται για την πρόβλεψη εκτιμήσεων σε σημεία που συνοδεύονται από ένα διάστημα, ενώ η διατακτική παλινδρόμηση (ordinal regression), τα δέντρα ταξινόμησης και παλινδρόμησης (classification and regression trees) και τα δίκτυα Bayes εφαρμόζονται για την εκτίμηση της προσπάθειας σε ένα προκαθορισμένο διάστημα.

[C2] MITTAS N., ANGELIS L. (2008). Partial Regression Error Characteristic Curves for the Comparison of Software Cost Prediction Models, *Workshop on Artificial Intelligence, Techniques in Software Engineering (AISEW 2008), July 2008, Patras, Greece.*

Η εργασία ασχολείται με τη σύγκριση μοντέλων κόστους λογισμικού. Προφανώς η σύγκριση μπορεί να γίνει με στατιστικούς ελέγχους αλλά αυτοί είναι καλό να συνοδεύονται και από γραφικές παραστάσεις που δείχνουν την κατανομή των σφαλμάτων με συγκριτικό τρόπο. Στην εργασία παρουσιάζονται οι μερικές χαρακτηριστικές καμπύλες σφάλματος παλινδρόμησης (PREC καμπύλες) οι οποίες χρησιμεύουν για την οπτικοποίηση της αθροιστικής κατανομής των σφαλμάτων πρόβλεψης σε περιορισμένα πεδία τιμών ώστε να εντοπιστούν τάσεις στις επιδόσεις των μοντέλων που συγκρίνονται. Η εργασία αυτή αποτελεί ουσιαστικά μια επέκταση της C3.

[C3] MITTAS N., ANGELIS L. (2008). Comparing Software Cost Prediction Models by a Visualization Tool. *34th EUROMICRO Conference on Software Engineering and Advanced Applications (SEAA), September 2008, Parma, Italy.*

Στην εργασία αυτή ουσιαστικά γίνεται για πρώτη φορά η εισαγωγή των καμπύλων REC (Regression Error Characteristic curves) στην περιοχή της εκτίμησης κόστους. Το βασικό πρόβλημα είναι η σύγκριση διαφόρων μοντέλων που κατασκευάζονται πάνω στο ίδιο σύνολο δεδομένων. Η ακρίβεια πρόβλεψης, ουσιαστικά μετριέται με τον υπολογισμό συναρτήσεων σφάλματος. Οι καμπύλες REC μπορούν να παραστήσουν με εύκολο τρόπο την κατανομή των σφαλμάτων πρόβλεψης αρκετών μοντέλων και να δείξουν με βάση τη θέση των καμπυλών ποιο μοντέλο υπερισχύει. Οι καμπύλες αυτές μπορούν να γίνουν εύκολα κατανοητές και από μη-ειδικούς και έχουν ενδιαφέρουσες ιδιότητες. Στην εργασία υπάρχει εφαρμογή της μεθοδολογίας σε δύο γνωστά σύνολα δεδομένων.

[C4] MITTAS N., ANGELIS L. (2008). Combining Regression and Estimation by Analogy in a Semi-parametric Model for Software Cost Estimation. *International Symposium on Empirical Software Engineering and Measurement ESEM'08, October 9–10, 2008, Kaiserslautern, Germany.*

Δύο από τις πιο γνωστές και δημοφιλείς μεθόδους εκτίμησης κόστους λογισμικού στη βιβλιογραφία είναι η ανάλυση παλινδρόμησης και η εκτίμηση με αναλογίες. Τα αποτελέσματα από πολλές εμπειρικές μελέτες δείχνουν ότι υπάρχει ασυμφωνία στην εξαγωγή πορίσματος σχετικά με το ποια είναι η καλύτερη από τις δύο μεθόδους από τις οποίες η παλινδρόμηση είναι παραμετρική, ενώ η άλλη μη-παραμετρική. Στην εργασία αυτή παρουσιάζουμε ένα ημι-παραμετρικό μοντέλο που κατορθώνει να ενσωματώσει παραμετρική πληροφορία σε ένα μη-παραμετρικό μοντέλο, συνδυάζοντας με αυτόν τον τρόπο την παλινδρόμηση και τη μέθοδο των αναλογιών. Επιπρόσθετα, δείχνουμε τη διαδικασία κατασκευής του μοντέλου αυτού σε δύο γνωστά σύνολα δεδομένων και παρουσιάζουμε την ακρίβεια πρόβλεψής του χρησιμοποιώντας αρκετά μέτρα ακρίβειας και συγκρίνοντας με την παλινδρόμηση και τη μέθοδο των αναλογιών όταν αυτές εφαρμόζονται μόνες τους. Τα αποτελέσματα δείχνουν ότι το ημι-παραμετρικό μοντέλο δίνει πιο ακριβείς προβλέψεις από τα άλλα δύο.

[C5] MITTAS N., L. ANGELIS (2009). Bootstrap Confidence Intervals for Regression Error Characteristic Curves Evaluating the Prediction Error of Software Cost Estimation Models. *2nd Artificial Intelligence Techniques in Software Engineering Workshop (AISEW2009), at the 5th IFIP*

Conference on Artificial Intelligence Applications and Innovations, April, Thessaloniki, Greece. Proceedings online, <http://ceur-ws.org/Vol-475>, pp. 221-230.

Η εργασία αυτή επεκτείνει τις προηγούμενες πάνω στις REC καμπύλες, παρουσιάζοντας μια μεθοδολογία υπολογισμού διαστημάτων εμπιστοσύνης για όλη την καμπύλη. Λόγω της ασυμμετρίας στην κατανομή των σφαλμάτων, υπάρχει αδυναμία υποθέσεων θεωρητικών κατανομών και για το λόγο αυτόν χρησιμοποιούμε μια bootstrap μεθοδολογία προσομοίωσης, η οποία μας επιτρέπει να υπολογίζουμε διαστήματα εμπιστοσύνης για τις REC καμπύλες. Με την κατασκευή των διαστημάτων μπορούμε να εξάγουμε στατιστικά συμπεράσματα για το αν είναι σημαντική η διαφορά της ικανότητας πρόβλεψης δύο μοντέλων.

[C6] MITTAS N., ANGELIS L. (2009). Bootstrap Prediction Intervals for a Semi-Parametric Software Cost Estimation Model. 35th Euromicro Conference on Software Engineering and Advanced Applications (SEAA 2009). August 2009, Patras, Greece, Proceedings published by IEEE, pp. 293-299.

Η εργασία αυτή επεκτείνει τις εργασίες J4 και C4 όπου ορίστηκε ένα ημι-παραμετρικό μοντέλο εκτίμησης κόστους λογισμικού συνδυάζοντας την παλινδρόμηση ελαχίστων τετραγώνων (Least Squares-LS) και την εκτίμηση με αναλογίες (Estimation by Analogy-EbA) για την κατασκευή ενός μοντέλου που ονομάζεται LSEbA. Στην εργασία χρησιμοποιούμε τη μέθοδο bootstrap για να υπολογίσουμε διαστήματα πρόβλεψης. Τα διαστήματα που παίρνουμε από την LSEbA συγκρίνονται με αυτά των άλλων μεθόδων με τη βοήθεια ενός νέου μέτρου σύγκρισης διαστημάτων που προτείνουμε και παίρνει υπόψη όχι μόνο την ικανότητα ορθής πρόβλεψης του πραγματικού κόστους αλλά και την ομοιότητα και το εύρος τους.

[C7] MITTAS N., ARGYROPOYLOY V. ANGELIS L. (2010). Modeling the Relationship between Software Effort and Size Using Deming Regression. To be presented at 6th International Conference on Predictive Models in Software Engineering. September 12-13, 2010, Timisoara, Romania.

Η σχέση μεταξύ της προσπάθειας (effort) που καταβάλλεται για την ανάπτυξη ενός έργου λογισμικού και του μεγέθους (size) του έργου έχει μοντελοποιηθεί μέχρι σήμερα ως εκθετική. Πιο αναλυτικά, ο λογάριθμος της προσπάθειας εκφράζεται ως μία γραμμική σχέση με τον λογάριθμο του μεγέθους. Η πιο γνωστή προσέγγιση που χρησιμοποιείται για την εκτίμηση των παραμέτρων του γραμμικού μοντέλου είναι η παλινδρόμηση ελαχίστων τετραγώνων (Least Squares-LS). Η συγκεκριμένη μέθοδος λαμβάνει υπόψη μόνο το σφάλμα της εξαρτημένης μεταβλητής ενώ οι μετρήσεις της ανεξάρτητης μεταβλητής θεωρείται ότι λαμβάνονται χωρίς σφάλμα. Η βασική ιδέα της συγκεκριμένης εργασίας είναι ότι στην πράξη η παραπάνω υπόθεση δεν είναι αληθής, από τη στιγμή που η μέτρηση του μεγέθους ενός έργου λογισμικού βασίζεται στην ακρίβεια του εργαλείου μέτρησης του κώδικα και συχνά σε υποκειμενικές κρίσεις του τεχνολόγου λογισμικού. Στην παρούσα εργασία εξετάζεται μία τεχνική παλινδρόμησης (Deming Regression-DR), η οποία λαμβάνει υπόψη το σφάλμα της μέτρησης της ανεξάρτητης μεταβλητής. Η DR εφαρμόζεται σε τέσσερα σύνολα δεδομένων και τα αποτελέσματα της συγκρίνονται με εκείνα της παλινδρόμησης ελαχίστων τετραγώνων. Το τελικό συμπέρασμα που προκύπτει από τα μέτρα ακρίβειας και τους στατιστικούς ελέγχους υποθέσεων είναι ότι η DR αποτελεί μία κατάλληλη τεχνική για τα σύνολα δεδομένων κόστους λογισμικούς καθώς λαμβάνει υπόψη την αβεβαιότητα της μέτρησης του μεγέθους του έργου.

[C8] KOSTI M., MITTAS N., ANGELIS L. (2010). DD-EbA: An algorithm for determining the number of neighbors in cost estimation by analogy using distance distributions. Workshop on Artificial Intelligence, Techniques in Software Engineering (AISEW 2010), September 2010, Ayia Napa, Cyprus.

Η συλλογιστική βασισμένη σε περιπτώσεις (Case Based Reasoning-CBR) και ειδικότερα η εκτίμηση με αναλογίες (Estimation by Analogy-EbA) έχει χρησιμοποιηθεί σε έναν μεγάλο αριθμό προβλημάτων,

όπως στην εκτίμηση κόστους λογισμικού (Software Cost Estimation-SCE). Οι συμβατικές μέθοδοι, εκτός από την έλλειψη ενός κριτηρίου για την επιλογή των κοντινότερων γειτονικών έργων, βασίζουν την εκτίμηση τους σε ένα προκαθορισμένο και σταθερό αριθμό γειτόνων από ολόκληρο το σύνολο ιστορικών έργων. Η παραπάνω προσέγγιση θέτει συγκεκριμένους περιορισμούς στην ικανότητα εκτίμησης των παραπάνω αλγορίθμων εξαιτίας του γεγονότος ότι δε λαμβάνουν υπόψη τη μοναδικότητα του κάθε υπό εκτίμηση έργου. Σε αυτή την εργασία προτείνεται μία νέα μεθοδολογία η DD-EbA, η οποία χρησιμοποιεί την έννοια των κατανομών των αποστάσεων μαζί με τη μετρική της απόστασης σε κάθε περίπτωση του συνόλου δεδομένων. Τα αποτελέσματα φανερώουν ότι η προτεινόμενη τεχνική έχει μικρότερο υπολογιστικό κόστος καθώς και την ίδια απόδοση πρόβλεψης με τη συμβατική EbA.

[C9] HAMDAN K., ANGELIS L. MITTAS N. (2010). Estimating learning by analogy: A case study in the UAE university. *International Conference of Education, Research and Innovation (ICERI 2010), November 15-17, 2010, Madrid, Spain.*

Η εργασία αυτή έχει στόχο να μετρήσει και σημαντικότερα, να προβλέψει την “ποσότητα” της εκμάθησης σε “problem-based” προγράμματα εκμάθησης. Εξαιτίας των δυσκολιών που υπάρχουν στην ποσοτικοποίηση θεμάτων εκπαίδευσης όπως της εκμάθησης, η οποία είναι μία αφηρημένη έννοια όταν πρόκειται να μετρηθεί, το πρόβλημα της ποσοτικοποίησης δεν είναι ένα τυπικό πρόβλημα το οποίο μπορεί κάποιος να εξετάσει με κλασικές μεθόδους στατιστικής ανάλυσης. Η διατάξιμη φύση των δεδομένων, οι οποίες λαμβάνονται από τις αξιολογήσεις απαιτεί κατάλληλες μεθόδους εκτίμησης. Στην παρούσα εργασία, γίνεται χρήση μίας μεθοδολογίας, η οποία είναι γνωστή ως μέθοδος των αναλογιών (Estimation by Analogy-EbA), γνωστή από την εφαρμογή της σε προβλήματα της τεχνολογίας λογισμικού (Software Engineering-SE). Πιο συγκεκριμένα, γίνεται χρήση της μεθόδου ώστε να εκτιμηθεί η ποσότητα εκμάθησης σε μία τάξη και βασίζεται σε παλαιότερα δεδομένα, τα οποία περιέχουν πληροφορία από παρόμοιες μετρήσεις. Η μεθοδολογία είναι γενικά απλή και δεν απαιτείται ισχυρό μαθηματικό υπόβαθρο και θεωρήσεις, καθώς βασίζεται στη φιλοσοφία της εύρεσης της κοντινότερης περίπτωσης από ένα σύνολο παλαιότερων περιπτώσεων. Η μέθοδος εφαρμόστηκε σε ένα σύνολο από 56 τάξεις Μαθηματικών και Επιστήμης Υπολογιστών, το οποίο συλλέχθηκε στο τέλος του ακαδημαϊκού έτους 2005 στο πανεπιστήμιο UAE των Ηνωμένων Αραβικών Εμιράτων.

[C10] MITTAS N. (2011) Evaluating the Performances of Software Cost Estimation Models through Prediction Intervals. *International Conference on Econophysics, June 2-3, 2011, Kavala, Greece.*

Η διαδικασία εκτίμησης με ακρίβεια του κόστους που απαιτείται για την ολοκλήρωση ενός καινούργιου έργου λογισμικού είναι μία πολύ σημαντική δραστηριότητα στο ερευνητικό πεδίο της εκτίμησης κόστους λογισμικού (Software Cost Estimation-SCE), καθώς σχετίζεται άμεσα με τις υποχρεώσεις ενός διοικητή έργου και τις αποφάσεις που καλείται να πάρει ώστε ο οργανισμός που αναλαμβάνει την ολοκλήρωση του έργου να μπορεί να κλείσει νέα συμβόλαια, να προβεί στη σωστή διαχείριση των έργων καθώς και να προκαθορίσει τον προϋπολογισμό ενός νέου συστήματος. Παρόλα αυτά, η ακριβής εκτίμηση του κόστους πραγματοποιείται με μεγάλη αβεβαιότητα και για τους λόγους αυτούς υπάρχει μία ασυμφωνία στα πειραματικά αποτελέσματα. Ο κύριος λόγος μπορεί να πηγάζει από ένα έμφυτο χαρακτηριστικό των μεθόδων εκτίμησης καθώς παράγουν προβλέψεις χωρίς να λαμβάνουν υπόψη το ρίσκο που υπάρχει σε ολόκληρη τη διαδικασία εκτίμησης. Στην παρούσα μελέτη, προτείνεται ένα στατιστικό πλαίσιο εργασίας που επικεντρώνεται στην κατασκευή διαστημάτων εμπιστοσύνης (Prediction Intervals-PI), τα οποία με τη σειρά τους παρέχουν με μία οπτιμιστική και πεσιμιστική εκτίμηση του πραγματικού κόστους. Το προτεινόμενο πλαίσιο ενσωματώνει διαφορετικούς δείκτες ακρίβειας, τυπικούς στατιστικούς ελέγχους υποθέσεων και γραφική απεικόνιση της ικανότητας πρόβλεψης ενώ εφαρμόζεται ακόμα σε ένα πραγματικό σύνολο δεδομένων με παλαιότερα έργα λογισμικού.

[G1] ΜΗΤΤΑΣ Ν., ΑΓΓΕΛΗΣ Ε. (2006). Διαστήματα Εμπιστοσύνης και Έλεγχοι Υποθέσεων στη Σύγκριση Μεθόδων Εκτίμησης Κόστους Λογισμικού. *Πρακτικά 19ου Πανελληνίου Συνεδρίου Στατιστικής, Καστοριά*

Η παράδοση ποιοτικού λογισμικού μέσα στα προκαθορισμένα χρονικά όρια και τα προβλεπόμενα χρονοδιαγράμματα είναι ο σημαντικότερος στόχος πολλών οργανισμών σε όλο τον κόσμο. Η εκτίμηση του κόστους ανάπτυξης λογισμικού με μεγάλη ακρίβεια είναι μία από τις σπουδαιότερες δραστηριότητες και για το λόγο αυτό έχουν προταθεί διάφορες τεχνικές εκτίμησης, οι οποίες ποικίλουν από αυθαίρετες, εμπειρικές προσεγγίσεις μέχρι τεκμηριωμένες μαθηματικά τεχνικές. Γενικά, υπάρχει μεγάλη συζήτηση για την επιλογή της καλύτερης τεχνικής που πρέπει να εφαρμόζει ένας οργανισμός. Στην πράξη, η επιλογή αυτή βασίζεται κυρίως σε κάποια στατιστικά μέτρα ακρίβειας που υπολογίζονται από τα σφάλματα των εκτιμήσεων. Η συνήθης πρακτική όμως της απλής σύγκρισης των τιμών τους χωρίς περαιτέρω ελέγχους υποθέσεων συχνά προκαλεί παραπλανητικά συμπεράσματα. Σε αυτή την εργασία, προτείνουμε τον συστηματικό έλεγχο, με χρήση στατιστικής συμπερασματολογίας, για τον καθορισμό της ανωτερότητας μιας μεθόδου σε σχέση με κάποια άλλη. Οι μέθοδοι που χρησιμοποιούμε βασίζονται σε τεχνικές κατασκευής διαστημάτων εμπιστοσύνης μέσω επαναληπτικής δειγματοληψίας (bootstrap) και μεταθέσεων (permutation tests).

[G2] ΔΙΜΟΚΑΣ Ν., ΜΙΤΤΑΣ Ν., ΝΑΝΟΠΟΥΛΟΣ Α., ΑΝΓΕΛΙΣ Λ. (2008). A Prototype System for Educational Data Warehousing and Mining. *12th Pan-Hellenic Conference on Informatics PCI 2008 August 2008, Samos, Proceedings published by IEEE.*

Τα πανεπιστήμια αντιμετωπίζουν συνεχώς αυξανόμενες απαιτήσεις από νομοθέτες και κοινότητες που αναζητούν έγκυρες πληροφορίες για τις επιδόσεις των φοιτητών και την αποδοτικότητα του πανεπιστημιακού συστήματος. Οι πληροφορίες αυτές δεν αφορούν μόνο τα πανεπιστήμια αλλά και την κυβέρνηση η οποία χρηματοδοτεί τις σπουδές. Η εργασία αυτή αφορά την ανάπτυξη μιας πρωτότυπης εκπαιδευτικής αποθήκης δεδομένων που αναπτύχθηκε στο Τμήμα Πληροφορικής του ΑΠΘ και η οποία υποστηρίζει την ανάλυση των εκπαιδευτικών δεδομένων. Στην εργασία παρουσιάζεται ο σχεδιασμός και η ανάπτυξη της αποθήκης δεδομένων και προτείνεται μια σειρά από στατιστικές μεθόδους και μεθόδους εξόρυξης δεδομένων για την εξαγωγή χρήσιμων συμπερασμάτων από τα εκπαιδευτικά δεδομένα της βάσης.

[G3] ΜΗΤΤΑΣ Ν., ΑΓΓΕΛΗΣ Ε. (2009). Γραφική Σύγκριση Μοντέλων Εκτίμησης Κόστους Λογισμικού Με Χαρακτηριστική Ανάλυση Σφάλματος Παλινδρόμησης. *Πρακτικά 22ου Πανελληνίου Συνεδρίου Στατιστικής, Χανιά.*

Στην εργασία αυτή γίνεται μια συνολική περιγραφή της στατιστικής μεθοδολογίας που βασίζεται στις Χαρακτηριστικές Καμπύλες Σφάλματος Παλινδρόμησης (Regression Error Characteristic-REC Curves). Η REC ανάλυση παρουσιάζεται προσαρμοσμένη ειδικά στο πρόβλημα της εκτίμησης κόστους λογισμικού και σε αυτή τη βάση προτείνεται ένα γενικότερο πλαίσιο συγκριτικής αξιολόγησης των μοντέλων πρόβλεψης και λήψης αποφάσεων ως προς το καταλληλότερο μοντέλο.

[G4] ΜΗΤΤΑΣ Ν., ΦΛΩΡΟΥ Γ., ΠΟΛΥΧΡΟΝΙΔΟΥ Π. (2009). Διάρκεια Φοίτησης στο Τμήμα Λογιστικής ΤΕΙ Καβάλας μέσω της Ανάλυσης Επιβίωσης. *Πρακτικά 22ου Πανελληνίου Συνεδρίου Στατιστικής, Χανιά.*

Στην εργασία αυτή, γίνεται μία στατιστική ανάλυση που αφορά στο τμήμα Λογιστικής του ΤΕΙ Καβάλας. Μέσω της περιγραφικής στατιστικής ανάλυσης, καθώς και διάφορων τεχνικών πολυμεταβλητής ανάλυσης αναζητούμε συσχετίσεις μεταξύ των παραγόντων που επηρεάζουν τη

διάρκεια φοίτησης και το βαθμό πτυχίου των αποφοίτων σπουδαστών. Θέλοντας να λάβουμε υπ' όψη μας όλη τη διαθέσιμη πληροφορία, ακόμα και από τους μη-πτυχιούχους σπουδαστές, χρησιμοποιούμε την ανάλυση επιβίωσης (Survival Analysis), μία στατιστική μεθοδολογία η οποία εφαρμόζεται ευρέως στη Βιοστατιστική.

[G5] POLYCHRONIDOU P., MITTAS N., FLOROU G. (2009). Factors that Affect the Duration of Studies of the Accounting Department of the Technological Institution of Kavala. 5th Pan-Hellenic Data Analysis Conference with International Participation, September 2009, Rethimno.

Τα τριτοβάθμια ιδρύματα της Ελλάδας πρέπει να πληροφορούν και να ενημερώνουν την επιστημονική κοινότητα και τα σχετικά υπουργεία για την πρόοδο των φοιτητών τους καθώς και για τη γενικότερη εικόνα των ιδρυμάτων. Στην εργασία αυτή, γίνεται προσπάθεια να αναλύσουμε με στατιστικό τρόπο τη διάρκεια φοίτησης στο τμήμα Λογιστικής του ΤΕΙ Καβάλας. Αρχικά, περιγράφουμε διάφορα δημογραφικά χαρακτηριστικά των φοιτητών του τμήματος κατά το πέρασμα των χρόνων ενώ με διάφορες τεχνικές πολυμεταβλητής ανάλυσης προσπαθούμε να ανιχνεύσουμε παράγοντες που επηρεάζουν τη διάρκεια φοίτησης και το βαθμό πτυχίου των τελειόφοιτων.

[G6] ΣΑΛΤΑΣ Β., ΚΑΛΑΜΠΑΚΑΣ Α., ΚΟΓΚΕΤΣΩΦ Α., ΠΟΛΥΧΡΟΝΙΔΟΥ Π., ΜΗΤΤΑΣ Ν., ΤΣΙΑΝΤΟΣ Β. (2010). Αξιολόγηση του Επιπέδου των Μαθηματικών των Πρωτοετών Φοιτητών της Σχολής Τεχνολογικών Εφαρμογών του ΤΕΙ Καβάλας. Πρακτικά 23ου Πανελληνίου Συνεδρίου Στατιστικής, Βέροια.

Η έρευνα στα μαθηματικά μπορεί να οριστεί ως σημαντική συγκέντρωση πληροφοριών για ένα ενδεχόμενο, με σκοπό να βοηθήσει ή να επικυρώσει επιστημονικές μελέτες. Το περιεχόμενο της επικεντρώνεται στη διδασκαλία και εμπέδωση των μαθηματικών γνώσεων, στον κοινωνικό ρόλο των μαθηματικών, στις διδακτικές καταστάσεις των μαθηματικών και στη σχέση των μαθηματικών με άλλες επιστήμες.

Η εν λόγω έρευνα μελετά τα αποτελέσματα της αξιολόγησης του επιπέδου των μαθηματικών των πρωτοετών φοιτητών της Σχολής Τεχνολογικών Εφαρμογών του ΤΕΙ Καβάλας. Χωρίζεται σε τρία βασικά στάδια αξιολόγησης: φοιτητή, εκπαιδευτικού συστήματος και καθηγητού. Έχει αρχικά ολοκληρωθεί η συγκέντρωση και η ανάλυση των αποτελεσμάτων σε ότι αφορά την αξιολόγηση των πρωτοετών φοιτητών και συγκεκριμένα στον προσδιορισμό του γνωστικού τους μαθηματικού επιπέδου σχετικά με τις γνώσεις και δεξιότητες που έχουν αποκτήσει ή όχι στη δευτεροβάθμια εκπαίδευσή τους. Θα ακολουθήσει και η αξιολόγησή τους και ως φοιτητές μεγαλύτερων ετών.

Χρησιμοποιήθηκε ένα ερωτηματολόγιο με 20 ερωτήσεις πολλαπλής επιλογής γενικών μαθηματικών γνώσεων δευτεροβάθμιας εκπαίδευσης. Ζητήθηκε να απαντήσουν οι πρωτοετείς φοιτητές των σχετικών τμημάτων. Οι συμμετέχοντες φοιτητές ομαδοποιήθηκαν σε τρεις βασικές κατηγορίες ανάλογα με το λύκειο και την κατεύθυνση: ΤΕΕ, Θετική, Τεχνολογική και Θεωρητική. Οι τέσσερις αυτές ομάδες χωρίστηκαν σε τρεις υποομάδες ανάλογα με το βαθμό των πανελληνίων στα μαθηματικά και συγκεκριμένα: 0-10, 11-15 και 16-20.

Η στατιστική μελέτη των απαντήσεων των ερωτηθέντων φοιτητών θα έχει ως αποτέλεσμα την πρόταση για αναδιοργάνωση ή όχι του τρόπου και του όγκου διδασκαλίας βασικών μαθηματικών εννοιών στη Σχολή Τεχνολογικών Εφαρμογών του Τεχνολογικού Εκπαιδευτικού Ιδρύματος Καβάλας.

[G7] ΤΣΙΑΝΤΟΣ Β., ΚΑΠΕΝΗΣ Κ., ΠΑΤΣΙΛΙΑΣ Γ., ΜΗΤΤΑΣ Ν., ΧΑΤΖΗΦΩΤΙΟΥ Σ. (2011). Η χρήση του Βίντεο στη Διδασκαλία των Μαθηματικών: Μία Πιλοτική Εφαρμογή στο Μάθημα «Μαθηματικά II» του Τμήματος Βιομηχανικής Πληροφορικής του ΤΕΙ Καβάλας. Πρακτικά 2ου Πανελληνίου Συνεδρίου Ένταξη των ΤΠΕ στην Εκπαιδευτική Διαδικασία, Μάρτιος 2011, Πάτρα.

Η εργασία αυτή αναλύει μία πιλοτική εφαρμογή βιντεοσκοπημένων διαλέξεων για την διδασκαλία των μαθηματικών στην τριτοβάθμια εκπαίδευση με σκοπό την διερεύνηση της χρησιμότητας του εργαλείου

αυτού προκειμένου να αποτελέσει συμπληρωματικό υλικό της εκπαιδευτικής διαδικασίας. Πιο συγκεκριμένα, βιντεοσκοπήθηκε το σύνολο των παραδόσεων της διδασκόμενης θεωρίας (ορισμών, μεθόδων, κλπ.) του μαθήματος “Μαθηματικά II” του Τμήματος Βιομηχανικής Πληροφορικής του ΤΕΙ Καβάλας. Στην συνέχεια έγινε επεξεργασία των βίντεο και χωρισμός τους σε βίντεο μικρότερης διάρκειας, περίπου των 6-20 λεπτών το καθένα, ταξινομημένα σε θεματικές ενότητες, τα οποία περιέχουν αναλυτική θεωρία και ένα παράδειγμα, αναλυτικά λυμένο, για κάθε μέθοδο. Τέλος, με τη χρήση ενός δομημένου ερωτηματολογίου, έγινε μία πρώτη αξιολόγηση του εργαλείου από είκοσι φοιτητές. Το ερευνητικό ερώτημα ήταν εάν τα βίντεο θα χρησίμευαν σαν βοήθημα στην μελέτη των φοιτητών, σε περιπτώσεις απουσίας από το μάθημα, ή σαν υπενθύμιση της διδασκαλίας κατά την επανάληψη των μεθόδων. Τα αποτελέσματα της έρευνας δείχνουν μία θετική στάση των φοιτητών απέναντι στην μέθοδο των βιντεοσκοπημένων διαλέξεων ως συμπληρωματικού υλικού διδασκαλίας.