



Πανεπιστήμιο Πειραιώς – Τμήμα Πληροφορικής
Πρόγραμμα Μεταπτυχιακών Σπουδών
«Πληροφορική»

Μεταπτυχιακή Διατριβή

Τίτλος Διατριβής	Στατιστική μελέτη του τρόπου πληκτρολόγησης και ανάπτυξη εφαρμογής για τη συλλογή δεδομένων
Όνοματεπώνυμο Φοιτητή	Ανδρέας Τσιγκρής
Πατρώνυμο	Ευάγγελος
Αριθμός Μητρώου	ΜΠΠΛ/ 09063
Επιβλέπων	Μαρία Βίρβου, Καθηγήτρια

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Τριμελής Εξεταστική Επιτροπή

(υπογραφή)

(υπογραφή)

(υπογραφή)

Όνομα Επώνυμο
Βαθμίδα

Όνομα Επώνυμο
Βαθμίδα

Όνομα Επώνυμο
Βαθμίδα

Ευχαριστίες

Με την ολοκλήρωση των σπουδών μου στο ΠΜΣ «Πληροφορική» του ΠΑ.ΠΕΙ θα ήθελα να ευχαριστήσω την επιβλέπουσα Καθηγήτρια μου Βίρβου Μαρία για τη εμπιστοσύνη που μου έδειξε αναθέτοντάς μου ένα τόσο ενδιαφέρον θέμα και για τις υποδείξεις που μου παρείχε κατά την εκπόνηση της παρούσας διατριβής.

Θα ήθελα ακόμα να ευχαριστήσω τον Δρ. Αλέπη Ευθύμιο για την καθοδήγηση και την αμέριστη συμπαράσταση που μου προσέφερε σε όλα τα στάδια αυτής της διατριβής.

Επίσης θα ήθελα να εκφράσω την εκτίμηση και τις ευχαριστίες μου σε όλο το διδακτικό προσωπικό του Προγράμματος Μεταπτυχιακών Σπουδών «Πληροφορική» του Πανεπιστημίου Πειραιά για την συνεργασία και τις γνώσεις τις οποίες μου προσέφεραν κατά τη διάρκεια της φοίτησής μου.

Ακόμη θα ήθελα να ευχαριστήσω τον Δ/ντη του 2^{ου} ΣΕΚ Αθηνών Κωνσταντίνου Αθανάσιο για την άδεια που μου παραχώρησε, να χρησιμοποιήσω το εργαστήριο Πληροφορικής του 1^{ου} ΕΠΑΛ Καισαριανής και τους Υπεύθυνους Εργαστηρίου, Εκπαιδευτικούς Μεγαλιό Νικόλαο και Πάνενα Ιωάννη χωρίς τη βοήθεια και υποστήριξη των οποίων δεν θα ήταν δυνατόν να πραγματοποιηθεί η έρευνα αυτής της διατριβής.

Τέλος, ένα μεγάλο ευχαριστώ στους μαθητές του 1ου ΕΠΑΛ Καισαριανής για την εθελοντική συμμετοχή τους σε αυτήν την έρευνα

Περίληψη

Ο τρόπος πληκτρολόγησης είναι ένα βιομετρικό γνώρισμα που τα θεμελιώδη χαρακτηριστικά του είναι ο χρόνος πίεσης των πλήκτρων και η χρόνος καθυστέρησης μεταξύ δύο διαδοχικών πιέσεων. Οι χρόνοι αυτοί επηρεάζονται από πλήθος παραγόντων που αφορούν είτε το περιβάλλον είτε την ψυχολογική και συναισθηματική κατάσταση του ατόμου. Οι χρονικές καθυστερήσεις όμως διαφορετικών διγραμμάτων διαφέρουν όταν οι παραπάνω παράγοντες παραμένουν σταθεροί.

Σκοπός της παρούσας διατριβής είναι να μελετηθούν οι παράγοντες που διαφοροποιούν το χρόνο καθυστέρησης διαφορετικών διγραμμάτων όταν οι συνθήκες του περιβάλλοντος και η συναισθηματική και ψυχολογική κατάσταση του ατόμου παραμένουν σταθερές.

Για το σκοπό λοιπόν αυτό σχεδιάστηκε και υλοποιήθηκε εφαρμογή για τη συλλογή των δεδομένων πληκτρολόγησης των χρηστών και τον κατάλληλο μετασχηματισμό τους για την μετέπειτα στατιστική επεξεργασία τους.

Η εφαρμογή εγκαταστάθηκε στους υπολογιστές του εργαστηρίου πληροφορικής του 1^{ου} ΕΠΑ.Λ Καισαριανής και χρησιμοποιήθηκε από 9 μαθητές του συγκεκριμένου σχολείου που συμμετείχαν εθελοντικά στην έρευνα.

Τα αποτελέσματα της έρευνας-μελέτης που πραγματοποιήθηκε στα πλαίσια αυτής της μεταπτυχιακής διατριβής οδήγησαν σε ορισμένα πολύ χρήσιμα συμπεράσματα, αλλά ταυτόχρονα γέννησαν νέα και πιο ενδιαφέροντα ερωτήματα.

Abstract

Keystroke dynamics is a biometric feature whose fundamental characteristics are the duration of each keystroke and the latency between two successive keystrokes. These characteristics are affected by many factors relating to either the environment or the psychological and emotional state of the individual. The keystrokes latencies between different digraphs differ when these factors remain constant.

The purpose of this thesis is to study the factors that modulate the keystroke latency different digraphs when the environmental conditions and the emotional and psychological state of the individual remain constant.

For this purpose we designed and implemented application for the collection of users typing data and the appropriate transformation for the subsequent statistical processing

The application was installed on the computers of 1st ΕΠΑ.Λ Kaisarianis' computer lab and used by 9 students of this school who voluntarily participated in the survey.

The results of the survey-study conducted in this master thesis led to some very useful information, but also gave birth to new and more interesting questions.

ΠΕΡΙΕΧΟΜΕΝΑ

Κεφάλαιο 1°	7
Εισαγωγή	7
1.1 Εισαγωγικό σημείωμα	7
1.2 Σκοπός της εργασίας	7
1.3 Διάρθρωση της εργασίας	7
Κεφάλαιο 2°	9
Βιομετρικά χαρακτηριστικά	9
2.1 Γενικά	9
2.2 Ποιοτικά γνωρίσματα βιομετρικών χαρακτηριστικών	9
2.3 Αναγνώριση – Αυθεντικοποίηση	9
2.4 Βιομετρικές μέθοδοι	10
2.5 Δείκτες απόδοσης	10
2.6 Τρόπος πληκτρολόγησης	10
2.7 Πλεονεκτήματα και μειονεκτήματα τρόπου πληκτρολόγησης	11
2.8 Ανασκόπηση της σχετικής βιβλιογραφίας	11
Κεφάλαιο 3°	13
Περιγραφή ανάπτυξης προγράμματος	13
3.1 Απαιτήσεις εφαρμογής	13
3.1.1 Σχεδίαση Βάσης Δεδομένων	13
3.2 Ανάλυση-Σχεδιασμός	15
3.2.1 Καθορισμός Απαιτήσεων	15
3.2.2 Καθορισμός σχέσεων μεταξύ κλάσεων	15
3.2.3 Μοντέλο Περιπτώσεων Χρήσης	15
Περιπτώσεις χρήσης	15
3.2.3.1 Εγγραφή χρήστη	17
3.2.3.2 Έλεγχος εισόδου χρήστη στο περιβάλλον της εφαρμογής	19
3.2.3.3 Πραγματοποίηση του τεστ από τον χρήστη	21
3.2.3.4 Πληκτρολόγηση λέξεων από τον χρήστη	23
3.2.3.5 Ρύθμιση παραμέτρων σύνδεσης με βάση δεδομένων	25
3.2.3.6 Εξαγωγή δεδομένων από την βάση δεδομένων	27
3.2.4 Διάγραμμα Κλάσεων	28
Κεφάλαιο 4°	29
Περιγραφή προγράμματος	29
4.1 Ρυθμίσεις σύνδεσης με τη βάση δεδομένων	29
4.2 Εγγραφή χρήστη	30
4.3 Είσοδος χρήστη	32
4.4 Πραγματοποίηση τεστ	33
4.5 Πληκτρολόγηση λέξεων	36
4.6 Είσοδος διαχειριστή	37
Κεφάλαιο 5°	38
Περιγραφή έρευνας	38
Γενικά	38
5.1 Προεργασία	39
5.2 Οι λέξεις του τεστ	39
5.3 Διεξαγωγή έρευνας	39
5.4 Εθελοντές	40
5.5 Συλλογή δεδομένων	40
5.6 Στατιστική επεξεργασία δεδομένων	40

Κεφάλαιο 6°	41
6.1 Στατιστικοί έλεγχοι και αποτελέσματα	41
6.2 Αυθεντικοποίηση χρήστη με τον τρόπο πληκτρολόγησης	50
6.3 Αναγνώριση χρήστη με τον τρόπο πληκτρολόγησης	52
6.4 Συμπεράσματα - Περίληψη	53
Βιβλιογραφία	54

Κεφάλαιο 1

Εισαγωγή

1.1 Εισαγωγικό σημείωμα

Ένα από τα σημαντικότερα τεχνολογικά επιτεύγματα του 2^{ου} μισού του 20^{ου} αιώνα είναι αναμφισβήτητα η εμφάνιση των ηλεκτρονικών υπολογιστών. Σήμερα οι υπολογιστές αποτελούν πλέον αναπόσπαστο κομμάτι κάθε πτυχής της καθημερινής μας ζωής. Οι υπολογιστές ελέγχουν σήμερα τις τηλεπικοινωνίες, τις αερομεταφορές καθώς και όλες τις οικονομικές συναλλαγές. Τους εμπιστευόμαστε από τα πιο απλά πράγματα, όπως οι αγορές μας μέσω διαδικτύου, μέχρι και κρίσιμες πληροφορίες σχετικά με τις τραπεζικές μας καταθέσεις και τα ιατρικά μας δεδομένα. Η ανάπτυξη της τεχνητής νοημοσύνης και της αλληλεπίδρασης ανθρώπου – υπολογιστή έχουν καταστήσει τους υπολογιστές απαραίτητο εργαλείο στην ενημέρωση, την εκπαίδευση και την απόκτηση γνώσης από την απέραντη δεξαμενή πληροφορίας που βρίσκεται στο διαδίκτυο.

Η χρήση υπολογιστών έχει απλοποιήσει και διευκολύνει κάθε ανθρώπινη δραστηριότητα και έχει ανοίξει νέους ορίζοντες σε όλους τους τομείς δραστηριοποίησης. Το τίμημα αυτού του γεγονότος είναι ολοένα και αυξανόμενη εξάρτηση της κοινωνίας μας από αυτούς. Αυτή η εξάρτηση απαιτεί, για ευνόητους λόγους την βελτίωση και την ανάπτυξη νέων μεθόδων στους τομείς της ασφάλειας και της αλληλεπίδρασης ανθρώπου υπολογιστή.

Εμείς στην εργασία μας αυτή θα ασχοληθούμε με τη μελέτη του τρόπου πληκτρολόγησης των χρηστών. Ο τρόπος πληκτρολόγησης είναι ένα ψυχολογικό βιομετρικό χαρακτηριστικό που χρησιμοποιείται για την αυθεντικοποίηση-αναγνώριση του χρήστη αλλά μπορεί να χρησιμοποιηθεί και για την καλύτερη επικοινωνία του ανθρώπου με τον υπολογιστή μιας και προσφέρει στον υπολογιστή πληροφορία που αφορά την τρέχουσα συναισθηματική, ψυχολογική και πνευματική κατάσταση του κάθε χρήστη.

1.2 Σκοπός της εργασίας

Σκοπός της εργασίας είναι η μελέτη και η κατανόηση του τρόπου πληκτρολόγησης των χρηστών και των παραγόντων από τους οποίους αυτός επηρεάζεται. Απαιτείται η δημιουργία προγράμματος που θα συλλέγει δεδομένα πληκτρολόγησης των χρηστών και θα τα μετασχηματίζει σε κατάλληλη μορφή ώστε να ακολουθήσει η στατιστική επεξεργασία τους.

1.3 Διάρθρωση της εργασίας

Στο επόμενο κεφάλαιο γίνεται εκτενής αναφορά των βιομετρικών χαρακτηριστικών. Αναφέρονται οι τρόποι ταξινόμησης τους και τα ποιοτικά γνωρίσματα τους καθώς και οι δείκτες μέτρησης της απόδοσης τους. Στη συνέχεια αναλύεται περαιτέρω το βιομετρικό χαρακτηριστικό που θα μελετήσουμε στην παρούσα διπλωματική εργασία, δηλαδή ο τρόπος πληκτρολόγησης. Παρουσιάζονται τα χαρακτηριστικά του καθώς και πλεονεκτήματα και τα μειονεκτήματα του.

Στο 3^ο κεφάλαιο γίνεται περιγραφή της ανάλυσης των απαιτήσεων και του σχεδιασμού των επιμέρους τμημάτων της εφαρμογής που θα αναπτυχθεί. Αναλύονται οι περιπτώσεις χρήσης της εφαρμογής συνοδευόμενες από διαγράμματα UML.

Στο 4^ο κεφάλαιο γίνεται παρουσίαση της υλοποιημένης εφαρμογής. Αναλύεται ο τρόπος χρήσης της και παρουσιάζονται οι περιπτώσεις χρήσης που περιγράφηκαν στο 3^ο κεφάλαιο.

Στο 5^ο κεφάλαιο αναλύεται το θεωρητικό υπόβαθρο της έρευνας που πραγματοποιήθηκε και περιγράφονται τα στάδια μέχρι την υλοποίηση της.

Τέλος, στο 6^ο κεφάλαιο παρουσιάζονται τα αποτελέσματα της στατιστικής επεξεργασίας των δεδομένων που προέκυψαν και προτείνονται τρόποι αξιοποίησης του τρόπου πληκτρολόγησης καθώς και προτάσεις βελτίωσης της απόδοσης του ως βιομετρικό χαρακτηριστικό.

Κεφάλαιο 2

Βιομετρικά χαρακτηριστικά

2.1 Γενικά

Τα βιομετρικά χαρακτηριστικά είναι μοναδικά, μετρήσιμα, φυσικά γνωρίσματα κάθε ατόμου που χρησιμοποιούνται προκειμένου αναγνωριστεί ή να επιβεβαιωθεί η ταυτότητα του. Τα βιομετρικά χαρακτηριστικά ταξινομούνται σε στατικά και δυναμικά και σε φυσιολογικά και ψυχολογικά (στοιχεία συμπεριφοράς). Τα στατικά-φυσιολογικά βιομετρικά παρουσιάζουν μεγάλη χρονική σταθερότητα και αποτελούν στοιχεία της φυσιολογίας του ατόμου όπως είναι πχ η ίριδα του ματιού, τα δακτυλικά αποτυπώματα κλπ. Αντίθετα, τα δυναμικά-ψυχολογικά βιομετρικά χαρακτηριστικά αφορούν στοιχεία ενεργειών του ατόμου και δεν είναι ιδιαίτερα σταθερά στο χρόνο. Παράγοντες όπως η κόυραση, το άγχος, η ψυχολογική διάθεση, η συναισθηματική κατάσταση του ατόμου παίζουν σημαντικότα ρόλο κατά την καταγραφή ενός ψυχολογικού βιομετρικού χαρακτηριστικού του.

2.2 Ποιοτικά γνωρίσματα βιομετρικών χαρακτηριστικών

- Καθολικότητα (universality): Αν το έχουν όλοι άνθρωποι.
- Διακριτότητα (distinctiveness): Αν μπορούν οι άνθρωποι να διακριθούν με βάση αυτό το χαρακτηριστικό.
- Μονιμότητα (permanence): Πόσο μόνιμο είναι το χαρακτηριστικό.
- Συλλεξιμότητα (collectability): Πόσο καλά μπορεί να αποκτηθεί και να ποσοτικοποιηθεί το βιομετρικό.

2.3 Αναγνώριση – Αυθεντικοποίηση

Αναγνώριση είναι η διαδικασία αναζήτησης της ταυτότητας ενός χρήστη μέσα στο πλήθος των εγγεγραμμένων χρηστών. Το σύστημα συγκρίνει τα στοιχεία με τα αποθηκευμένα στοιχεία των χρηστών στην προσπάθεια του να εντοπίσει το χρήστη.

Αυθεντικοποίηση είναι η διαδικασία επιβεβαίωσης της ταυτότητας του χρήστη. Το σύστημα συγκρίνει τα στοιχεία του χρήστη με τα αποθηκευμένα στοιχεία του. Είναι σαφώς πιο γρήγορη διαδικασία από την αναγνώριση.

Οι μέθοδοι που χρησιμοποιούνται, είτε μόνες τους είτε σε συνδυασμό μεταξύ τους, τόσο για την αναγνώριση όσο και για την αυθεντικοποίηση βασίζονται σε :

- κάτι που γνωρίζει ο χρήστης (π.χ. passwords, PINs κλπ.)
- κάτι που κατέχει ο χρήστης (π.χ. δελτία ταυτότητας, μαγνητικές κάρτες κλπ.)
- βιομετρικά χαρακτηριστικά του χρήστη (βιομετρικές μέθοδοι)

Οι δύο πρώτες κατηγορίες μεθόδων αποτελούν παραδοσιακές τεχνικές που χρησιμοποιούνται κατά κόρον σήμερα παρουσιάζοντας όμως σημαντικά μειονεκτήματα καθώς τα στοιχεία που απαιτούνται για την αναγνώριση – αυθεντικοποίηση μπορεί να χαθούν, να υποκλαπούν ή και να πλαστογραφηθούν.

Το πλεονέκτημα των βιομετρικών μεθόδων είναι ότι τα στοιχεία με βάση τα οποία επιτυγχάνονται η αναγνώριση και η αυθεντικοποίηση είναι χαρακτηριστικά γνωρίσματα κάθε ατόμου που το διαφοροποιούν από τα υπόλοιπα άτομα. Κατά συνέπεια τα στοιχεία αυτά είναι σχεδόν αδύνατο να χαθούν, να ξεχαστούν ή να αντιγραφούν.

2.4 Βιομετρικές μέθοδοι

Οι κυριότερες βιομετρικές μέθοδοι που χρησιμοποιούνται σήμερα είναι:

α) Στατικές – Φυσιολογικές μέθοδοι

- Σάρωση της ίριδας του ματιού
- Σάρωση του αμφιβληστροειδή
- Αναγνώριση προσώπου
- Δακτυλικά αποτυπώματα
- Γεωμετρία χεριού

β) Δυναμικές – Ψυχολογικές μέθοδοι

- Αναγνώριση φωνής
- Εξακρίβωση υπογραφής
- Αναγνώριση τρόπου πληκτρολόγησης

2.5 Δείκτες απόδοσης

Οι πιο συχνά χρησιμοποιούμενοι δείκτες της απόδοσης των βιομετρικών μεθόδων είναι:

- Το ποσοστό των λανθασμένων απορρίψεων, FRR (False Rejection Rate), δηλαδή το ποσοστό των εξουσιοδοτημένων χρηστών στους οποίους δε παραχωρήθηκε πρόσβαση
- Το ποσοστό των λανθασμένων αποδοχών, FAR (False Acceptance Rate), δηλαδή το ποσοστό των μη εξουσιοδοτημένων χρηστών στους οποίους παραχωρήθηκε πρόσβαση.

Οι δύο αυτοί δείκτες είναι αμοιβαίως αποκλειόμενοι, δηλαδή όταν αυξάνει ο ένας μειώνεται ο άλλος. Όπως είναι φυσικό, όταν το σύστημα γίνει πολύ αυστηρό ώστε να εξαλειφθεί το ποσοστό των λανθασμένων αποδοχών, θα δυσκολεύονται και οι πραγματικοί χρήστες να αποκτήσουν πρόσβαση σε αυτό με αποτέλεσμα να αυξάνει το ποσοστό των λανθασμένων απορρίψεων. Ακριβώς το αντίθετο συμβαίνει όταν το σύστημα γίνει αρκετά χαλαρό ώστε να αποκτούν εύκολα πρόσβαση όλοι οι 'νόμιμοι' χρήστες. Έγκειται στην κάθε εφαρμογή ξεχωριστά, αφού ελαχιστοποιήσει τους δύο αυτούς δείκτες, να θέσει τον επιθυμητό χαρακτήρα της εφαρμογής. Αν δηλαδή μεγαλύτερη έμφαση θα δοθεί στην αποτροπή των εισβολέων ή στην φιλικότητα του συστήματος προς τους νόμιμους χρήστες του.

2.6 Τρόπος πληκτρολόγησης

Παρ' όλες τις εξελίξεις στον τομέα της πληροφορικής η πιο συχνά χρησιμοποιούμενη συσκευή εισόδου δεδομένων στον υπολογιστή παραμένει το πληκτρολόγιο. Το γεγονός αυτό καθ' αυτό αποτελεί σημαντικό πλεονέκτημα και σχεδόν επιβάλλει τη μελέτη του τρόπου πληκτρολόγησης καθώς τα αποτελέσματα και τα συμπεράσματα που προκύπτουν βρίσκουν εφαρμογή σε πολλούς και διαφορετικούς τομείς της επιστήμης της πληροφορικής, όπως είναι η ασφάλεια, η αλληλεπίδραση ανθρώπου-υπολογιστή, η εξ αποστάσεως εκπαίδευση και πολλοί άλλοι. Τα τελευταία χρόνια πολλές έρευνες έχουν γίνει στον τομέα κυρίως της ασφάλειας όπου ο τρόπος πληκτρολόγησης χρησιμοποιείται για την ταυτοποίηση των χρηστών είτε μόνος του είτε σε συνδυασμό με τη χρήση συνθηματικών.

Ο τρόπος πληκτρολόγησης, όπως προαναφέρθηκε, είναι ένα ψυχολογικό βιομετρικό χαρακτηριστικό με ορισμένα εγγενή στοιχεία που αποτελούν μειονεκτήματα όσον αφορά τη χρήση του στον τομέα της ασφάλειας. Τα καταγραφόμενα βιομετρικά πρότυπα κάθε ατόμου διαφέρουν σημαντικά μεταξύ τους με αποτέλεσμα να παρουσιάζεται μεγάλη ενδο-ατομική μεταβλητότητα. Οι διαφορές αυτές όμως είναι δυνατό να μετατραπούν σε πλεονεκτήματα για άλλους τομείς καθώς αυτά μπορούν να αποτελέσουν δείκτες εκτίμησης της ψυχολογικής ή συναισθηματικής κατάστασης του ατόμου.

Η ανάλυση και επεξεργασία του τρόπου πληκτρολόγησης γίνεται με βάση τα παρακάτω χαρακτηριστικά:

- Χρονική Διάρκεια: Διάρκεια πίεσης του κάθε πλήκτρου (ή «χρόνος πίεσης »)
- Χρονική Καθυστέρηση: Διάρκεια μεταξύ διαδοχικών πληκτρολογήσεων (ή «χρόνος πτήσης»)
- Δύναμη που ασκείται κατά την πίεση κάθε πλήκτρου (απαιτείται ειδικό πληκτρολόγιο)
- Χρόνος πληκτρολόγησης συγκεκριμένων λέξεων

2.7 Πλεονεκτήματα και μειονεκτήματα τρόπου πληκτρολόγησης

Το βασικό πλεονέκτημα του είναι ότι για τη συλλογή των δεδομένων χρησιμοποιείται το πληκτρολόγιο και δεν απαιτούνται εξειδικευμένες συσκευές εισόδου που στις περισσότερες περιπτώσεις είναι δαπανηρές.

Επίσης η συλλογή και επεξεργασία πρωτογενών δεδομένων δεν είναι παρεμβατική, καθώς ο χρήστης δεν χρειάζεται να κάνει τίποτα παραπάνω από αυτό που προτίθεται, δηλαδή να γράψει στο πληκτρολόγιο.

Με βάση τα ποιοτικά χαρακτηριστικά που αναφέρθηκαν στο προηγούμενο κεφάλαιο, ο τρόπος πληκτρολόγησης παρουσιάζει υστέρηση έναντι των άλλων βιομετρικών χαρακτηριστικών:

Είναι ψυχολογικό χαρακτηριστικό με όλα τα εγγενή μειονεκτήματα των χαρακτηριστικών της κατηγορίας του. Δηλαδή, επηρεάζεται από το άγχος ή την κούραση του χρήστη αλλά και από τις συνθήκες του περιβάλλοντος όπως είναι η διαφορετική θέση πληκτρολόγησης, το διαφορετικό πληκτρολόγιο, ή γλώσσα γραφής κλπ

Παρουσιάζει χαμηλό βαθμό καθολικότητας καθώς είναι ένα χαρακτηριστικό που δεν απαντάται σε όλους τους ανθρώπους αλλά μόνο σε όσους ασχολούνται συστηματικά με τους υπολογιστές.

Παρουσιάζει χρονική αστάθεια μιας και επηρεάζεται σε μεγάλο βαθμό από την εμπειρία που αποκτά ο χρήστης ως προϊόν της ενασχόλησης και εξοικείωσης του με τον υπολογιστή.

Τέλος όσον αφορά το κριτήριο της διακριτότητας, αυτό αποτελεί αντικείμενο μελέτης των περισσότερων ερευνών για το αν και κατά πόσο τα χαρακτηριστικά που προκύπτουν από τον τρόπο πληκτρολόγησης είναι ικανά να διαφοροποιήσουν τους ανθρώπους μεταξύ τους.

2.8 Ανασκόπηση της σχετικής βιβλιογραφίας

Μία από τις πρώτες εργασίες πραγματοποιήθηκε 1985 από τους Umphress και Williams με τίτλο "Identity verification through keyboard characteristics". Σε αυτή την έρευνα χρησιμοποιήθηκε ως κριτήριο αυθεντικοποίησης μόνο η χρονική καθυστέρηση μεταξύ διαδοχικών πληκτρολογήσεων. Το κείμενο αναφοράς αποτελούνταν από 1400 χαρακτήρες ενώ το κείμενο αναγνώρισης από 300. Συμμετείχαν στο πείραμα 17 άτομα με διαφορετικές ικανότητες πληκτρολόγησης. Η αυθεντικοποίηση επιτυγχανόταν με στατιστικό έλεγχο της μέσης τιμής του χρόνου καθυστέρησης μεταξύ των δύο κειμένων (t-tests). Το μέγιστο FAR που επιτεύχθηκε ήταν 6% αλλά ο όγκος των δεδομένων που απαιτούνταν ήταν ιδιαίτερα μεγάλος.

Το 1990 οι Joyce και Gupta [1] δημοσίευσαν την εργασία τους με τίτλο "Identity authentication based on keystroke latencies". Σε αυτή την έρευνα συμμετείχαν 33 χρήστες. Κάθε χρήστης πληκτρολόγούσε 8 φορές το όνομα, το επώνυμο, το username και το password του ώστε να δημιουργηθεί το προσωπικό του προφίλ. Η αυθεντικοποίηση βασιζόταν στον υπολογισμό της μέσης τιμής και της τυπικής απόκλισης του χρόνου πτήσης κάθε διγράμματος που περιεχόταν στις παραπάνω συμβολοσειρές. Στη συνέχεια κάθε χρήστης πραγματοποίησε 5 προσπάθειες εισόδου (συνολο=33*5=165 προσπάθειες). Τέλος, 6 χρήστες επιλέχτηκαν τυχαία και οι υπόλοιποι 27 πραγματοποίησαν 5 προσπάθειες εισόδου χρησιμοποιώντας τα στοιχεία κάθε ενός από τους επιλεγμένους 6 χρήστες (σύνολο

27*5*6=810 προσπάθειες). Τα αποτελέσματα που προέκυψαν ήταν FAR ίσο με 0.25% και FRR ίσο με 16.67%.

Το 1997 οι Monrose και Rubin παρουσίασαν την εργασία τους με τίτλο "Authentication via Keystroke Dynamics". Σε αυτή χρησιμοποιήθηκε όχι μόνο η χρονική καθυστέρηση μεταξύ διαδοχικών πληκτρολογήσεων αλλά και η χρονική διάρκεια πίεσης των πλήκτρων. Τα κείμενα αναφοράς και αναγνώρισης είτε ήταν ίδια είτε ήταν επιλογή των χρηστών. Για την αναγνώριση των χρηστών χρησιμοποιήθηκαν διάφοροι αλγόριθμοι ταξινόμησης και τα αποτελέσματα έδειξαν ότι όταν :

Τα κείμενα αναφοράς και αναγνώρισης ήταν τα ίδια και προκαθορισμένα το ποσοστό επιτυχούς αναγνώρισης ήταν 90,7%.

Τα κείμενα αναφοράς και αναγνώρισης ήταν ίδια αλλά το κείμενο αναφοράς ήταν επιλογή του χρήστη, το ποσοστό αναγνώρισης μειώθηκε στο 44,1%

Τα κείμενα αναφοράς και αναγνώρισης ήταν επιλογή του χρήστη αλλά και διαφορετικά μεταξύ τους το ποσοστό επιτυχίας μειώθηκε ακόμα περισσότερο στο 23%.

Το 1999 οι ίδιοι δημοσίευσαν την εργασία τους με τίτλο «Keystroke dynamics as a biometric for authentication». Και σε αυτή χρησιμοποιήθηκαν και οι χρόνοι πτήσης των διγραμμάτων αλλά και οι χρόνοι πίεσης των γραμμάτων. Σκοπός της εργασίας ήταν η αναγνώριση και για το σκοπό αυτό τα δεδομένα που συλλέχθηκαν από 63 εθελοντές χωρίστηκαν σε κείμενα αναφοράς και σε κείμενα αναγνώρισης. Χρησιμοποιήθηκαν διάφορα ευρετικά όπως η ευκλείδεια απόσταση, η συνάρτηση πυκνότητας πιθανότητας με στάθμιση ή χωρίς κλπ και τα ποσοστά επιτυχίας κυμαίνονταν από 83,22%-92,14%.

Μια ενδιαφέρουσα προσέγγιση στην αναγνώριση χρήστη παρουσιάστηκε το 2005 στην εργασία «Dealing with Different Languages and Old Profiles in Keystroke Analysis of Free Text» των Daniele Gunetti, Claudia Picardi, and Giancarlo Ruffo. Η μόνη πληροφορία που χρησιμοποιήθηκε ήταν ο χρόνος πτήσης των παραγόμενων διγραμμάτων. Μεταξύ δύο κειμένων υπολογίζεται η μέση τιμή του χρόνου πτήσης των κοινών διγραμμάτων τους και τα διγράμματα αυτά ταξινομούνται κατ' αύξουσα σειρά. Στην συνέχεια υπολογίζεται η απόκλιση των θέσεων που βρίσκεται κάθε δίγραμμα στα δυο διανύσματα διγραμμάτων. Το άθροισμα αυτών των αποκλίσεων προς την μέγιστη απόκλιση χρησιμοποιείται ως δείκτης ομοιότητας. Η μέγιστη απόκλιση είναι $N^2/2$ ή $(N-1)^2/2$ αν ο αριθμός N των διγραμμάτων είναι άρτιος ή περιττός αντίστοιχα. Η έρευνα έγινε σε 30 εθελοντές, τα κείμενα των οποίων συλλέχθηκαν σε διάρκεια 1,5 έτους. Μετά από αυτό το διάστημα τους ζητήθηκε να πληκτρολογήσουν 2 κείμενα, 1 στα Αγγλικά και 1 στα Ιταλικά. Τα ποσοστά επιτυχίας ήταν 90% και 93% αντίστοιχα για το Ιταλικό και το Αγγλικό κείμενο. Το μειονέκτημα βέβαια στην συγκεκριμένη μέθοδο ήταν ο μεγάλος όγκος δεδομένων που απαιτούνταν αλλά τα αποτελέσματα έδειξαν ότι τα χαρακτηριστικά πληκτρολόγησης παραμένουν σταθερά για μεγάλα χρονικά διαστήματα.

Το 2007 δημοσιεύτηκε μια εργασία με τίτλο «Keystroke Statistical Learning Model for Web Authentication» από τους Cheng-Huang Jiang, Shiuhyng Shieh, και Jen-Chien Liu. Στην έρευνα συμμετείχαν 58 εθελοντές. Αρχικά, το σύστημα απαιτούσε από κάθε χρήστη να πληκτρολογήσει 20 φορές το username και το password του ώστε να δημιουργηθούν τα δείγματα αναφοράς. Στη συνέχεια οι χρήστες πραγματοποίησαν 15 προσπάθειες εισόδου στο λογαριασμό τους ενώ 257 άλλοι ανώνυμοι χρήστες προσπάθησαν να εισέλθουν παράνομα σε αυτούς τους λογαριασμούς. Κάθε λογαριασμός δέχτηκε 44-82 προσπάθειες εισβολής. Στη συνέχεια η αυθεντικοποίηση επιτυγχανόταν με το υπολογισμό της πιθανότητας να ανήκουν οι συμβολοσειρές που δίνονταν στον νόμιμο χρήστη. Η πιθανότητα αυτή υπολογιζόταν με βάση τη μέση τιμή και την τυπική απόκλιση του χρόνου πτήσης των διγραμμάτων των παραπάνω συμβολοσειρών. Τα αποτελέσματα έδειξαν FAR και FRR ίσα με 2% περίπου για συνθηματικά τουλάχιστον 9 χαρακτήρων.

Κεφάλαιο 3

Περιγραφή ανάπτυξης προγράμματος

3.1 Απαιτήσεις εφαρμογής

Για να μελετήσουμε τους χρόνους πίεσης των γραμμάτων και τους χρόνους πτήσης των διγραμμάτων των λέξεων που πληκτρολογεί ο χρήστης θα δημιουργήσουμε μία εφαρμογή η οποία θα καταγράφει αυτούς τους χρόνους, θα τους αποθηκεύει στη βάση δεδομένων, και θα δίνει τη δυνατότητα στον διαχειριστή να εξάγει τα δεδομένα σε αρχείο κειμένου ή EXCEL. Τα αρχεία αυτά θα αποτελέσουν αρχεία εισόδου για την εφαρμογή SPSS για την μετέπειτα στατιστική επεξεργασία των δεδομένων. Ακόμη θα υπάρχει ένα τεστ τοποθέτησης των γραμμάτων στις θέσεις που βρίσκονται στο πληκτρολόγιο. Επειδή η εφαρμογή πρόκειται να χρησιμοποιηθεί σε τοπικό δίκτυο υπολογιστών και θα υπάρχει μια βάση δεδομένων στην οποία θα αποθηκεύονται τα δεδομένα όλων των χρηστών που χρησιμοποιούν την εφαρμογή.

Σύμφωνα με τις απαιτήσεις που περιγράφονται παραπάνω το σύστημα θα αποτελείται:

α. Τμήμα Δικτύου – Επικοινωνία

Η εφαρμογή θα σχεδιαστεί με βάση την αρχιτεκτονική client-server και συγκεκριμένα θα είναι εφαρμογή 2 επιπέδων (2-tier). Δηλαδή η εφαρμογή θα σχεδιαστεί έτσι ώστε να υπάρχει ένας κεντρικός υπολογιστής (server) που θα φιλοξενεί τη βάση δεδομένων και όλοι οι υπόλοιποι υπολογιστές θα επικοινωνούν με αυτόν σε ενσύρματο ή ασύρματο δίκτυο LAN.

β. Δεδομένα – Λειτουργία

Όπως έχουμε αναφέρει θα δημιουργηθεί μια Βάση Δεδομένων η οποία θα περιέχει όλα τα στοιχεία που απαιτούνται για την πραγματοποίηση των διαδικασιών που αναφέρθηκαν.

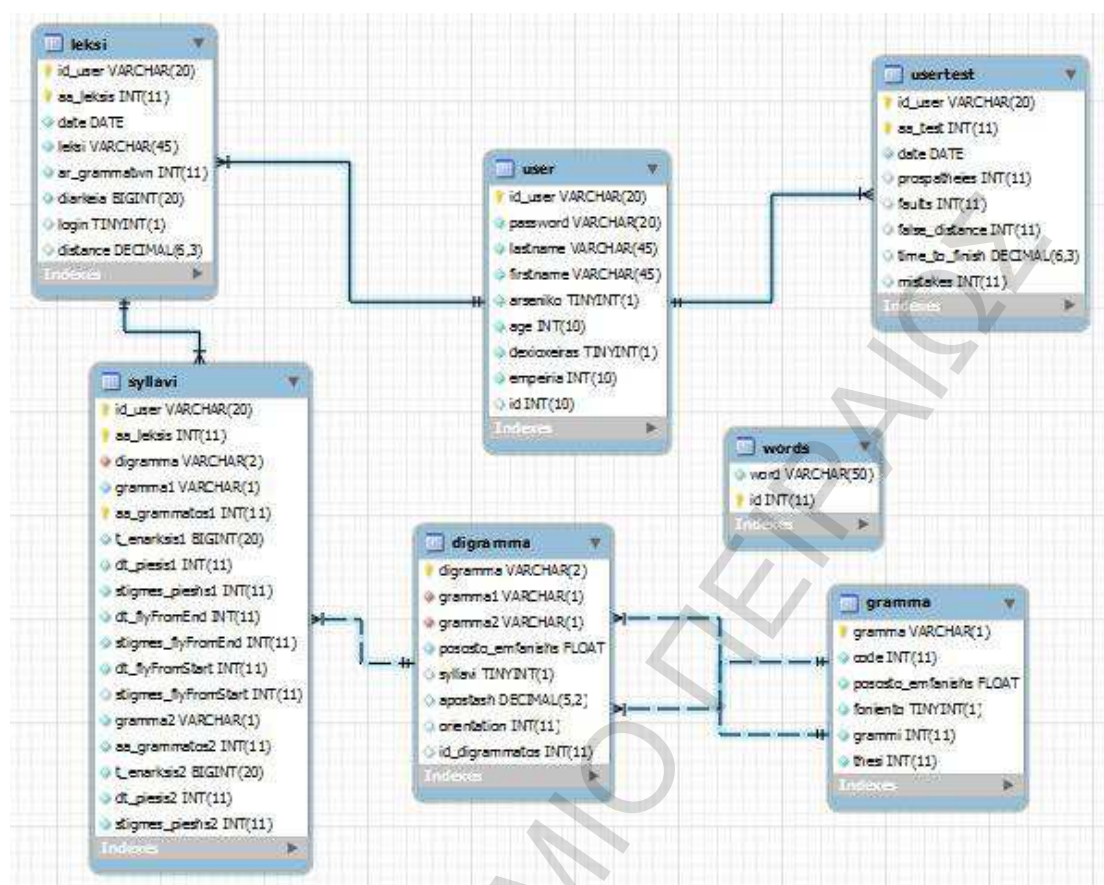
Ακόμα η Βάση Δεδομένων θα συμπεριλαμβάνει και όλους τους χρήστες που θα χρησιμοποιήσουν την εφαρμογή και εκτός από τα στοιχεία τους θα περιέχει το username και το password.

Η βάση δεδομένων θα αναπτυχθεί σύμφωνα με το σχεσιακό μοντέλο δεδομένων.

Το προτεινόμενο σύστημα διαχείρισης βάσης δεδομένων είναι η τελευταία έκδοση της ORACLE για δικτυακό περιβάλλον. Η Βάση Δεδομένων μας θα σχεδιαστεί με το πρόγραμμα MySQL Workbench 5.2 CE. Η λειτουργία του θα φιλοξενηθεί σε λειτουργικό σύστημα Windows XP ή Windows 7.

3.1.1 Σχεδίαση Βάσης Δεδομένων

Το διάγραμμα του σχεσιακού μοντέλου της βάσης δεδομένων που θα χρησιμοποιηθεί στην εφαρμογή που θα αναπτυχθεί είναι το παρακάτω:



Εικόνα 3.1 Διάγραμμα E-R της Βάσης Δεδομένων

Ο πίνακας user θα περιέχει τους χρήστες που συμμετέχουν στην έρευνα και στον οποίο θα αποθηκεύονται τα στοιχεία τους.

Ο πίνακας userTest θα περιέχει τα αποτελέσματα των τεστ που πραγματοποιούν οι χρήστες, δηλαδή θα αποθηκεύονται ο χρόνος ολοκλήρωσης του τεστ, οι προσπάθειες του χρήστη, τα λάθη αλλά και η απόσταση των λαθών. Ως απόσταση λάθους εννοείται η απόσταση της τοποθέτησης ενός γράμματος από το χρήστη από την ορθή θέση του πάνω στο πληκτρολόγιο.

Ο πίνακας leksi θα αποθηκεύει τις λέξεις που πληκτρολογεί ο κάθε χρήστης και στον οποίο θα περιέχονται τα χαρακτηριστικά της κάθε λέξης.

Ο πίνακας syllavi θα αποθηκεύει τις συλλαβές κάθε λέξης που πληκτρολογείται και στον οποίο θα περιέχονται οι χρόνοι πίεσης κάθε γράμματος της συλλαβής αλλά και ο χρόνος καθυστέρησης της συλλαβής.

Ο πίνακας digramma περιέχει όλα τα δυνατά διγράμματα της ελληνικής γλώσσας με τα χαρακτηριστικά τους όπως η απόσταση των γραμμάτων τους σε ένα QWERTY πληκτρολόγιο, η συχνότητα εμφάνισής τους στη γλώσσα και αν πρόκειται για συλλαβές της γλώσσας ή όχι.

Ο πίνακας gramma περιέχει τα 24 γράμματα της ελληνικής αλφαβήτου με τα χαρακτηριστικά τους, δηλαδή τα στατιστικά εμφάνισής τους στη γλώσσα, αν πρόκειται για φωνήεν ή σύμφωνο και την θέση του γράμματος σε ένα τυπικό πληκτρολόγιο.

Τέλος, ο πίνακας words περιέχει τις 52 λέξεις που καλούνται να πληκτρολογήσουν οι χρήστες κατά τη διάρκεια της έρευνας.

3.2 Ανάλυση-Σχεδιασμός

3.2.1 Καθορισμός Απαιτήσεων

Εξαγωγή αρχικής λίστας υποψηφίων κλάσεων

Από την περιγραφή απαιτήσεων υψηλού επιπέδου που δόθηκε, προκύπτει η ακόλουθη αρχική λίστα ουσιαστικών και πιθανών κλάσεων πεδίου προβλήματος (έχοντας μετατρέψει το πρόσωπο σε πρώτο ενικό):

Gramma	lexi
Digramma	administrator
user	UserTest
pliktro	PanelKeyboard
MySQLconnector	JavaToFile

Περιορισμός υποψηφίων κλάσεων

Η ανωτέρω λίστα μπορεί να περιοριστεί σημαντικά απαλείφοντας:

Αναφορές στο ίδιο το σύστημα λογισμικού που αναπτύσσουμε (Σύστημα διαχείρισης)

Αναφορές σε χειριστές του συστήματος που πρόκειται να αναπτύξουμε καθώς βρίσκονται "έξω" από τα όρια του συστήματος (χειριστής)

Αναφορές σε οντότητες που βρίσκονται εκτός του πεδίου του προβλήματος και κατά συνέπεια δεν πρόκειται να συμμετάσχουν ως οντότητες στην υλοποίηση των λειτουργιών του συστήματος (σχολείο).

3.2.2 Καθορισμός σχέσεων μεταξύ κλάσεων

Πρωταρχικός στόχος κατά την κατασκευή του μοντέλου του πεδίου προβλήματος είναι ο εντοπισμός σχέσεων μεταξύ των υποψηφίων κλάσεων. Ωστόσο, καθώς στην αντικειμενοστρεφή σχεδίαση υπάρχουν πολλά είδη σχέσεων μεταξύ κλάσεων, δεν κρίνεται σκόπιμο στο στάδιο αυτό να απεικονιστούν ιδιαίτερες λεπτομέρειες. Συνήθως αρκεί η απεικόνιση στο μοντέλο σχέσεων τύπου "έχει" (*has*) «αποτελείται από» και σχέσεων τύπου "είναι" (*is*). Το πρώτο είδος αναφέρεται σε σχέσεις περιεκτικότητας μεταξύ δύο κλάσεων που στην απλούστερη μορφή καλούνται συναρμολογήσεις (ή συσσωματώσεις). Μια συναρμολόγηση υποδηλώνει ότι αντικείμενα μιας κλάσης περιέχουν αντικείμενα μιας άλλης κλάσης. Η συναρμολόγηση απεικονίζεται ως μια γραμμή μεταξύ των κλάσεων με έναν λευκό ρόμβο στο άκρο της περιέχουσας κλάσης. Το δεύτερο είδος αναφέρεται σε σχέσεις κληρονομικότητας μεταξύ κλάσεων.

Η λέξη αποτελείται από δίγραμματα

Το δίγραμμα αποτελείται από γράμματα

3.2.3 Μοντέλο Περιπτώσεων Χρήσης (Use Case Model)

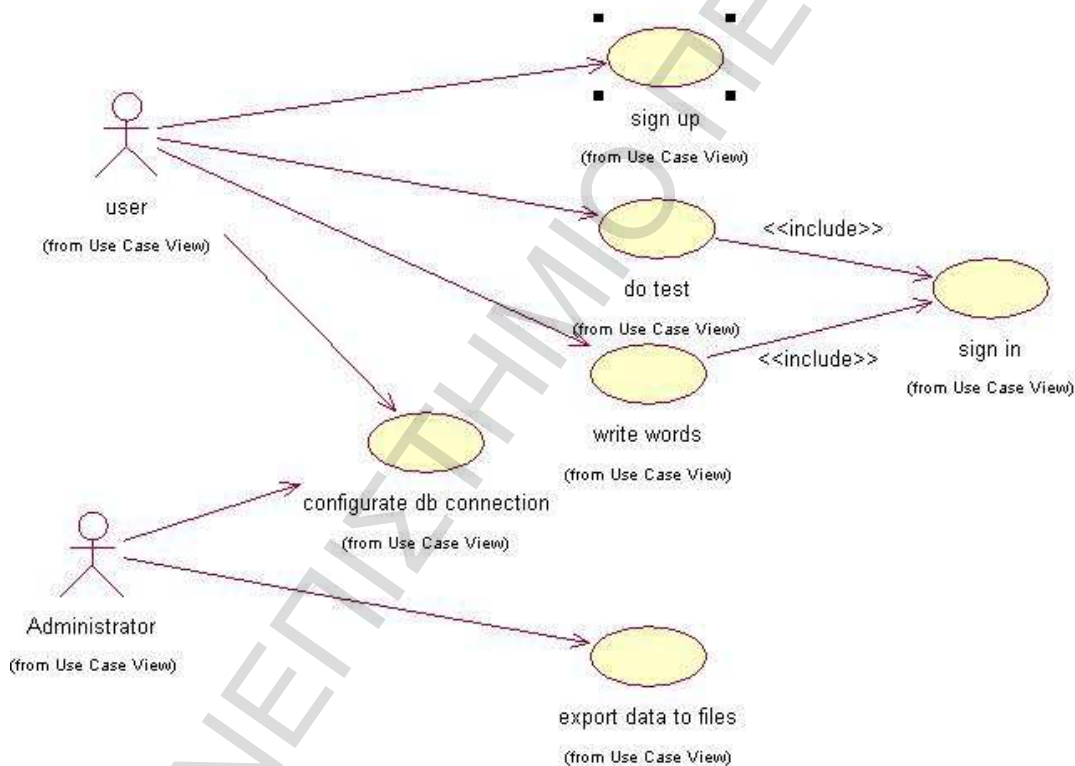
Περιπτώσεις χρήσης

Μία περίπτωση χρήσης (use case) είναι μια ακολουθία ενεργειών που ένας χρήστης του συστήματος (συνήθως πρόκειται για άνθρωπο αλλά μπορεί να είναι οποιαδήποτε εξωτερική οντότητα όπως ένα άλλο σύστημα) πραγματοποιεί στο σύστημα για να επιτύχει ένα συγκεκριμένο σκοπό.

Ο ορισμός μιας περίπτωσης χρήσης περιλαμβάνει όλες τις δυνατές συμπεριφορές, περιλαμβανομένης της κανονικής ή βασικής ροής (όπου για παράδειγμα εκτελούνται όλα όσα αναμένει ο χρήστης για την επίτευξη του στόχου) αλλά και όλων των "εναλλακτικών" ροών, όπου κάτι για παράδειγμα αποκλίνει από την "κανονική" συμπεριφορά. Μια περίπτωση χρήσης είναι επιτυχής όταν προδιαγράφει μόνο "ποια" είναι η συμπεριφορά του συστήματος χωρίς καμία αναφορά στο "πώς" αυτή επιτυγχάνεται, χωρίς δηλαδή περιγραφή τεχνικών λεπτομερειών υλοποίησης.

Με βάση τα παραπάνω προκύπτουν οι παρακάτω περιπτώσεις

- Έγγραφο χρήστη στο σύστημα
- Είσοδος χρήστη στο σύστημα
- Πραγματοποίηση τεστ
- Πληκτρολόγηση λέξεων
- Ρύθμιση παραμέτρων σύνδεσης με ΒΔ
- Εξαγωγή δεδομένων σε αρχείο xls και txt.



Εικόνα 3.2 Διάγραμμα περιπτώσεων χρήσης

3.2.3.1 Προδιαγραφή περίπτωσης χρήσης

1. Τίτλος περίπτωσης χρήσης

Εγγραφή χρήστη

2. Σύνομη περιγραφή

Ο χρήστης καταχωρεί τα στοιχεία του και εγγράφεται στον Κατάλογο Χρηστών.

3. Ροή γεγονότων

3.1 Βασική ροή

1. Ο Χειριστής «Χρήστης» από την ΑΡΧΙΚΗ ΟΘΟΝΗ επιλέγει την εργασία «ΕΓΓΡΑΦΗ ΝΕΟΥ ΧΡΗΣΤΗ»

2. Η εφαρμογή εμφανίζει την οθόνη «EggraphUserFrame» η οποία έχει πεδία συμπλήρωσης των στοιχείων που απαιτούνται και δύο κουμπιά «ΕΓΓΡΑΦΗ ΧΡΗΣΤΗ» και «ΑΚΥΡΩΣΗ»

3. Ο Χειριστής συμπληρώνει τα στοιχεία που περιέχονται στη φόρμα και πατάει το κουμπί «ΕΓΓΡΑΦΗ ΧΡΗΣΤΗ».

4. Η εφαρμογή ελέγχει την πληρότητα και εγκυρότητα των στοιχείων, ελέγχει αν το username υπάρχει και εισάγει μια νέα εγγραφή στον Κατάλογο Χρηστών

5. Ο έλεγχος επανέρχεται στο βήμα 1.

3.2 Εναλλακτικές ροές

3.2.1 Εναλλακτική ροή 1

3α. Ο χρήστης επιλέγει «ΕΠΙΣΤΡΟΦΗ».

4α. Η εφαρμογή κλείνει τη φόρμα και τερματίζει την εργασία.

3.2.2 Εναλλακτική ροή 2

3β. Τα στοιχεία που δόθηκαν είναι ελλιπή.

4β. Η εφαρμογή εμφανίζει μήνυμα στο χρήστη και επανέρχεται στο βήμα 2.

3.2.3 Εναλλακτική ροή 3

3β. Το username υπάρχει ήδη.

4β. Η εφαρμογή εμφανίζει μήνυμα στο χρήστη και επανέρχεται στο βήμα 2.

4. Μη λειτουργικές απαιτήσεις

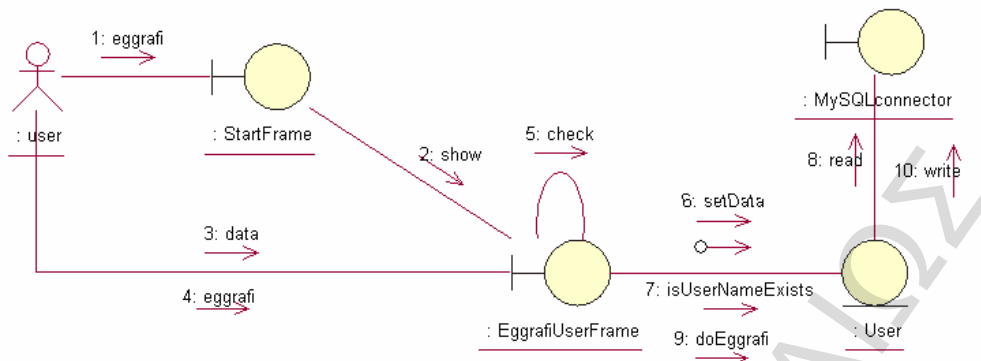
Δεν υπάρχουν για αυτή την περίπτωση χρήσης.

5. Κατάσταση εισόδου

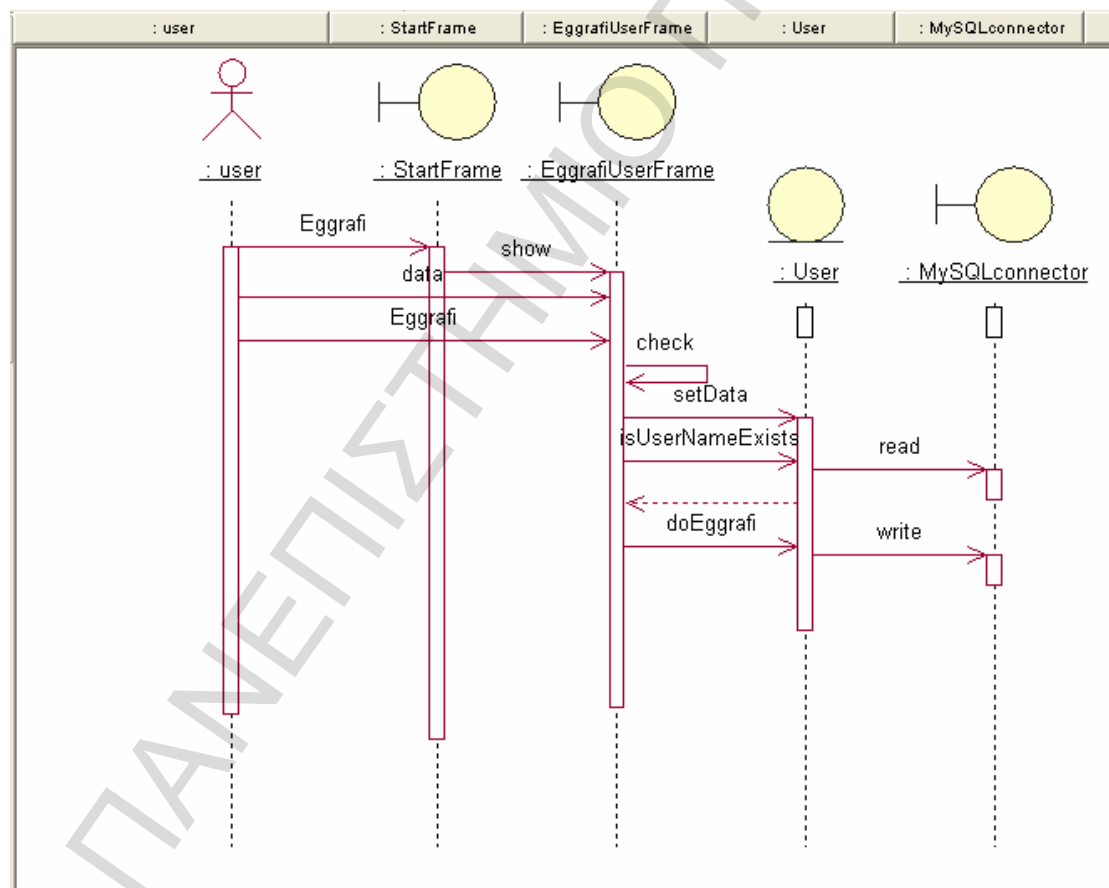
Δεν υπάρχουν ιδιαίτερες απαιτήσεις εισόδου στην περίπτωση χρήσης.

6. Κατάσταση εξόδου

Έχει προστεθεί 1 νέα εγγραφή στον Κατάλογο Χρηστών.



Εικόνα 3.3 Διάγραμμα ακολουθίας της περίπτωσης χρήσης 3.2.3.1



Εικόνα 3.4 Διάγραμμα συνεργασίας της περίπτωσης χρήσης 3.2.3.1

3.2.3.2 Προδιαγραφή περίπτωσης χρήσης

1. Τίτλος περίπτωσης χρήσης

Έλεγχος εισόδου χρήστη στο περιβάλλον της εφαρμογής

2. Σύντομη περιγραφή

Ο χρήστης συμπληρώνει τον κωδικό και το password ώστε να εισέλθει στο σύστημα.

3. Ροή γεγονότων

3.1 Βασική ροή

1. Ο χρήστης από την ΑΡΧΙΚΗ ΟΘΟΝΗ επιλέγει την εργασία «ΕΙΣΟΔΟΣ ΧΡΗΣΤΗ»

2. Η εφαρμογή εμφανίζει την οθόνη «AuthenticationFrame» η οποία έχει πεδία συμπλήρωσης για το username και password και τα κουμπιά «ΕΙΣΟΔΟΣ» και «ΕΠΙΣΤΡΟΦΗ»

3. Ο χρήστης συμπληρώνει τον κωδικό και το password και πατά το πλήκτρο «ΕΙΣΟΔΟΣ»

4. Ο εφαρμογή συνδέεται με τη βάση δεδομένων, επαληθεύει τα στοιχεία που δόθηκαν με τα στοιχεία του χρήστη, ανοίγει την ΚΥΡΙΑ ΟΘΟΝΗ του χρήστη.

3.2 Εναλλακτικές ροές

3.2.1 Εναλλακτική ροή 1

3α. Ο χρήστης επιλέγει «ΕΠΙΣΤΡΟΦΗ».

4α. Η εφαρμογή κλείνει τη φόρμα και τερματίζει την εργασία.

3.2.2 Εναλλακτική ροή 2

3β. Τα στοιχεία που δόθηκαν είναι ελλιπή.

4β. Η εφαρμογή εμφανίζει μήνυμα στο χρήστη και επανέρχεται στο βήμα 2.

3.2.2 Εναλλακτική ροή 3

3γ. Τα στοιχεία που δόθηκαν είναι λανθασμένα.

4γ. Η εφαρμογή εμφανίζει μήνυμα στο χρήστη και επανέρχεται στο βήμα 2.

4. Μη λειτουργικές απαιτήσεις

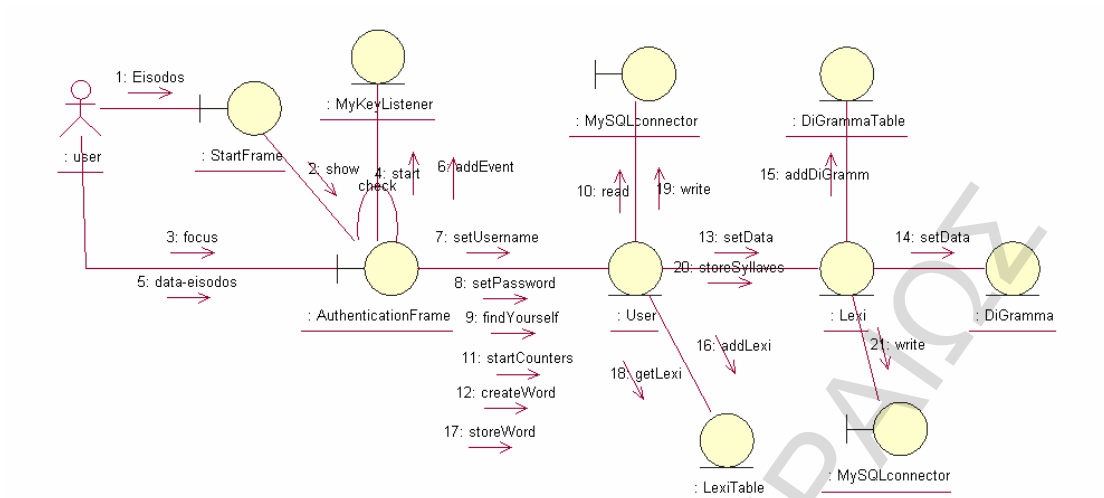
Δεν υπάρχουν για αυτή την περίπτωση χρήσης.

5. Κατάσταση εισόδου

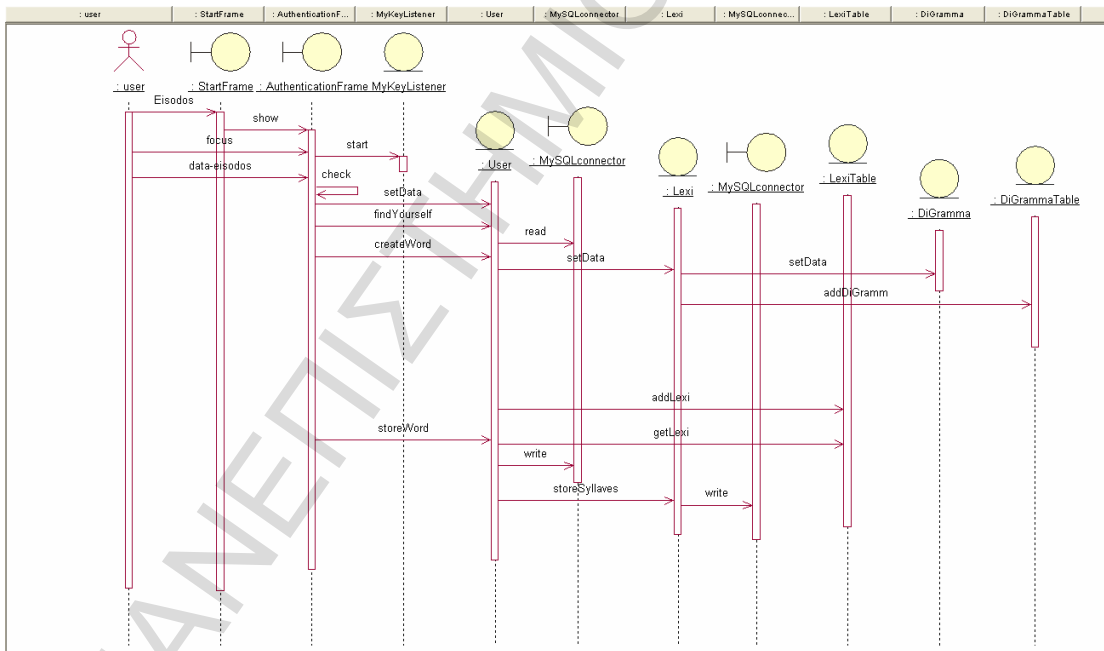
Δεν υπάρχουν ιδιαίτερες απαιτήσεις εισόδου στην περίπτωση χρήσης

6. Κατάσταση εξόδου

Εμφανίζεται η ΚΥΡΙΑ ΟΘΟΝΗ του χρήστη.



Εικόνα 3.5 Διάγραμμα ακολουθίας της περίπτωσης χρήσης 3.2.3.2



Εικόνα 3.6 Διάγραμμα συνεργασίας της περίπτωσης χρήσης 3.2.3.2

3.2.3.3 Προδιαγραφή περίπτωσης χρήσης

1. Τίτλος περίπτωσης χρήσης

Πραγματοποίηση του τεστ από τον χρήστη

2. Σύντομη περιγραφή

Ο χρήστης τοποθετεί τα γράμματα στις θέσεις που πιστεύει ότι βρίσκονται στο πληκτρολόγιο

3. Ροή γεγονότων

3.1 Βασική ροή

1. Ο χρήστης από την ΑΡΧΙΚΗ ΟΘΟΝΗ επιλέγει την εργασία «ΤΕΣΤ»

2. Η εφαρμογή εμφανίζει την οθόνη «PanelKeyboard» στην οποία απεικονίζεται ένα πληκτρολόγιο και τα 24 γράμματα του ελληνικού αλφαβήτου

3. Ο χρήστης με αριστερό κλικ του ποντικιού επιλέγει ένα από τα γράμματα και με drag and drop το τοποθετεί στην θέση που πιστεύει ότι βρίσκεται στο πληκτρολόγιο.

4. Η εφαρμογή χρωματίζει ανάλογα τη θέση και το γράμμα ώστε να φαίνονται ότι έχουν χρησιμοποιηθεί.

5. Ο έλεγχος επανέρχεται στο βήμα 2.

6. Μετά την τοποθέτηση όλων των γραμμάτων ο χρήστης πιέζει το κουμπί «ΔΕΙΤΕ ΤΑ ΑΠΟΤΕΛΕΣΜΑΤΑ».

7. Η εφαρμογή εμφανίζει μήνυμα με τα αποτελέσματα του χρήστη και τα αποθηκεύει στον ΚΑΤΑΛΟΓΟ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

3.2 Εναλλακτικές ροές

3.2.1 Εναλλακτική ροή 1

3α. Ο χρήστης κάνει διπλό αριστερό κλικ πάνω στη θέση που έχει τοποθετήσει προηγούμενα γράμμα.

4α. Η εφαρμογή αποδεσμεύει τη θέση και το γράμμα ώστε να είναι διαθέσιμα για νέα τοποθέτηση.

4. Μη λειτουργικές απαιτήσεις

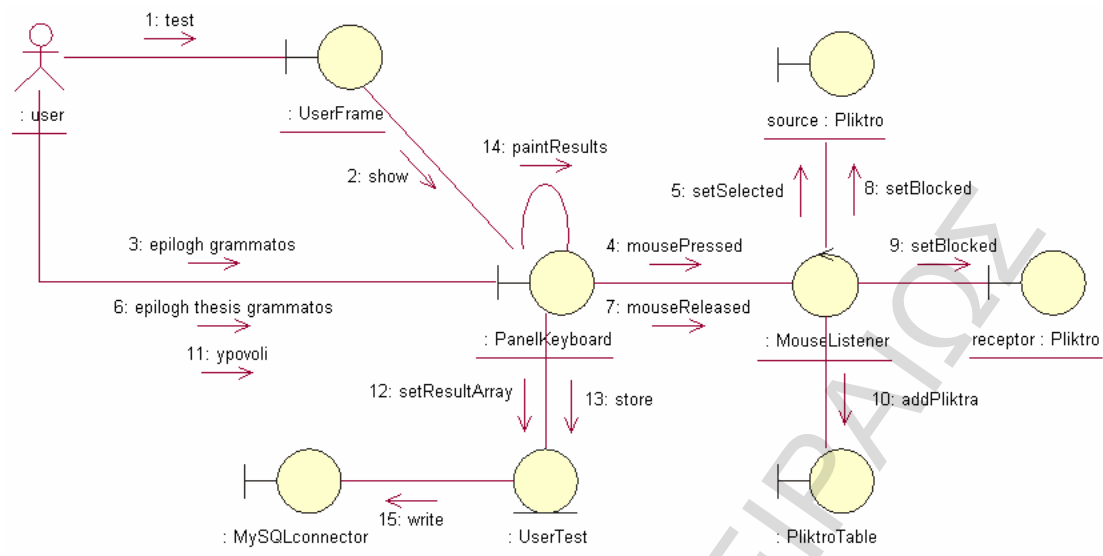
Δεν υπάρχουν για αυτή την περίπτωση χρήσης.

5. Κατάσταση εισόδου

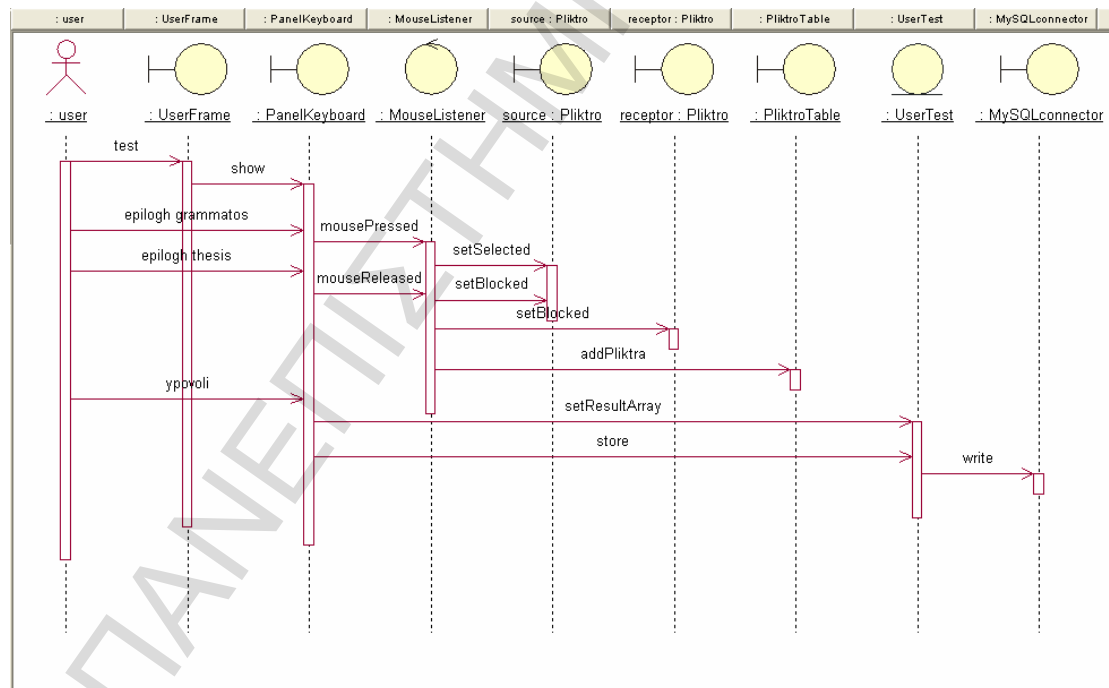
Δεν υπάρχουν ιδιαίτερες απαιτήσεις εισόδου στην περίπτωση χρήσης

6. Κατάσταση εξόδου

Εμφανίζεται η ΚΥΡΙΑ ΟΘΟΝΗ του χρήστη.



Εικόνα 3.7 Διάγραμμα ακολουθίας της περίπτωσης χρήσης 3.2.3.3



Εικόνα 3.8 Διάγραμμα συνεργασίας της περίπτωσης χρήσης 3.2.3.3

3.2.3.4 Προδιαγραφή περίπτωσης χρήσης

1. Τίτλος περίπτωσης χρήσης

Πληκτρολόγηση λέξεων από τον χρήστη

2. Σύντομη περιγραφή

Ο χρήστης πληκτρολογεί λέξεις που του παρουσιάζονται στην οθόνη της εφαρμογής

3. Ροή γεγονότων

3.1 Βασική ροή

1. Ο χρήστης από την ΚΥΡΙΑ ΟΘΟΝΗ επιλέγει την εργασία «ΠΛΗΚΤΡΟΛΟΓΗΣΗ»

2. Η εφαρμογή εμφανίζει την οθόνη «PanelRecord» στο πάνω μέρος της οποίας εμφανίζεται η λέξη που ο χρήστης καλείται να πληκτρολογήσει

3. Ο χρήστης πληκτρολογεί τη λέξη στο πλαίσιο που βρίσκεται στο κάτω μέρος της οθόνης

4. Η εφαρμογή ελέγχει κάθε γράμμα που πληκτρολογεί ο χρήστης ώστε να ταιριάζει με το αντίστοιχο γράμμα της προς αντιγραφή λέξης και το χρωματίζει πράσινο.

5. Ο χρήστης πιέζει SPACE ή ENTER

6. Η εφαρμογή καταχωρεί τη λέξη στον ΚΑΤΑΛΟΓΟ ΛΕΞΕΩΝ και εμφανίζει την επόμενη προς αντιγραφή λέξη

7. Ο έλεγχος επιστρέφει στο βήμα 2.

3.2 Εναλλακτικές ροές

3.2.1 Εναλλακτική ροή 1

3α. Ο χρήστης κάνει λάθος κατά την πληκτρολόγηση της λέξης

4α. Η εφαρμογή ειδοποιεί ηχητικά και οπτικά τον χρήστη και τον καθοδηγεί με κατάλληλο μήνυμα.

4. Μη λειτουργικές απαιτήσεις

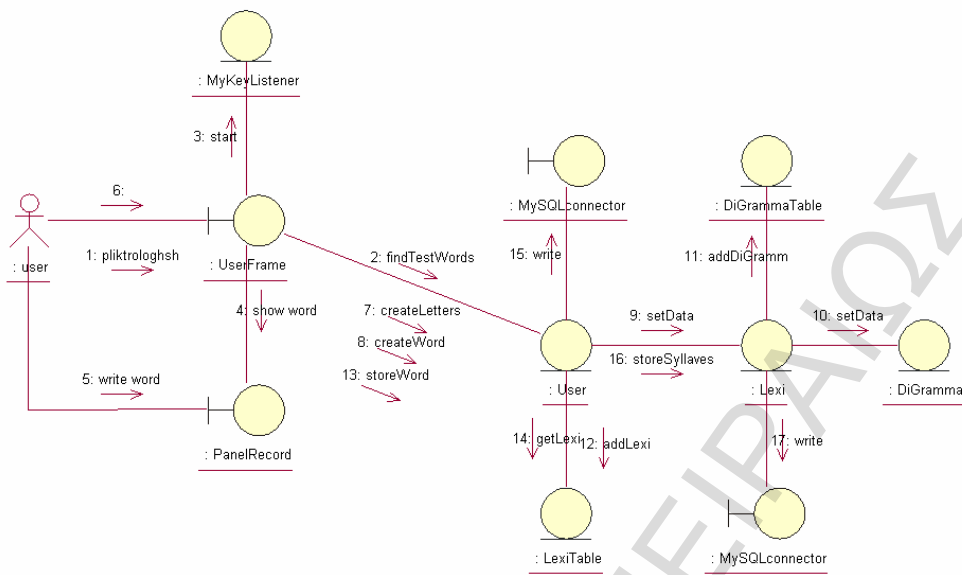
Δεν υπάρχουν για αυτή την περίπτωση χρήσης.

5. Κατάσταση εισόδου

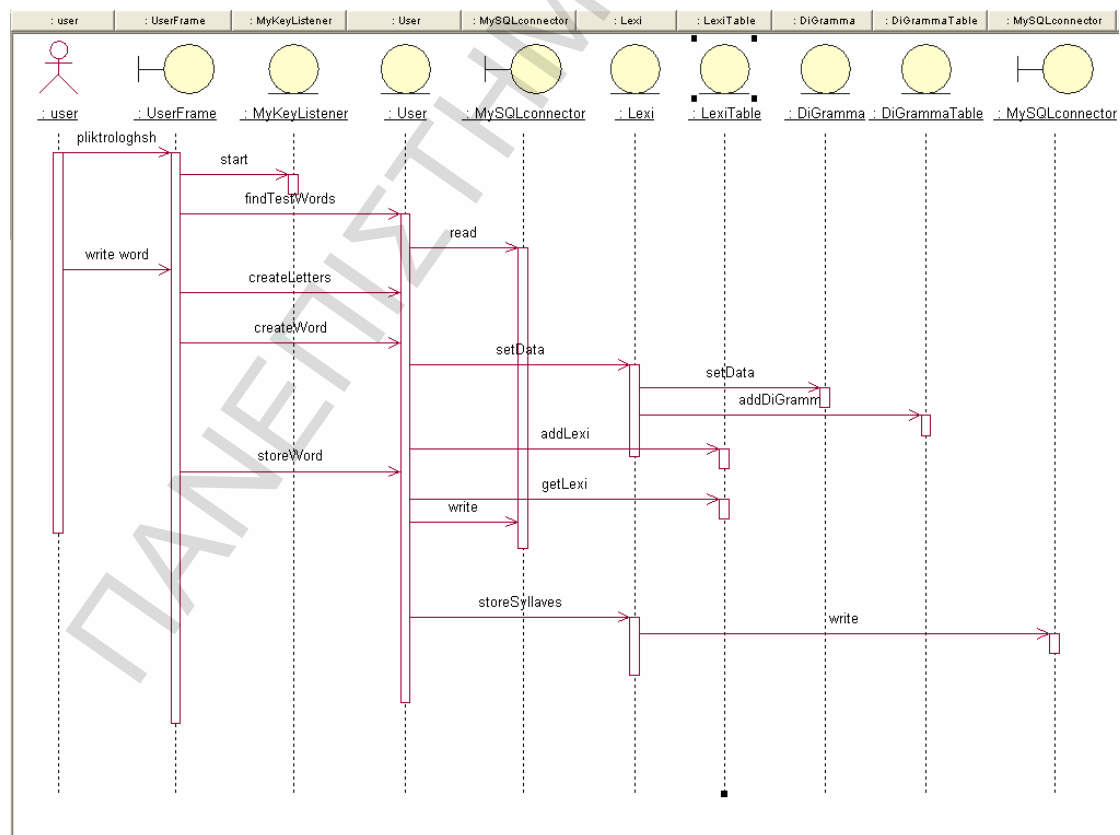
Δεν υπάρχουν ιδιαίτερες απαιτήσεις εισόδου στην περίπτωση χρήσης

6. Κατάσταση εξόδου

Εμφανίζεται η ΚΥΡΙΑ ΟΘΟΝΗ του χρήστη.



Εικόνα 3.9 Διάγραμμα ακολουθίας της περίπτωσης χρήσης 3.2.3.4



Εικόνα 3.10 Διάγραμμα συνεργασίας της περίπτωσης χρήσης 3.2.3.4

3.2.3.5 Προδιαγραφή περίπτωσης χρήσης

1. Τίτλος περίπτωσης χρήσης

Ρύθμιση παραμέτρων σύνδεσης με βάση δεδομένων

2. Σύντομη περιγραφή

Ο χρήστης ρυθμίζει τις παραμέτρους που απαιτούνται για την σύνδεση με τη ΒΔ

3. Ροή γεγονότων

3.1 Βασική ροή

1. ο χρήστης από την ΑΡΧΙΚΗ ΟΘΟΝΗ επιλέγει την εργασία «ΡΥΘΜΙΣΕΙΣ ΣΥΝΔΕΣΗΣ»

2. η εφαρμογή εμφανίζει την οθόνη «DBconfigFrame» .

3. Ο χρήστης επιλέγει τις παραμέτρους που επιθυμεί και πιέζει το πλήκτρο «TEST CONNECTION»

4. Η εφαρμογή ελέγχει τις παραμέτρους, επικοινωνεί με τον database server και εμφανίζει τις βάσεις δεδομένων

5. ο χρήστης επιλέγει τη βάση δεδομένων και πιέζει το πλήκτρο «ΣΥΝΔΕΣΗ»

6. Η εφαρμογή συνδέεται με τη βάση δεδομένων, αποθηκεύει τις παραμέτρους σύνδεσης σε αρχείο καταγραφής κλείνει τη φόρμα και τερματίζει την εργασία.

3.2 Εναλλακτικές ροές

3.2.1 Εναλλακτική ροή 1

3α. Τα στοιχεία που δόθηκαν είναι ελλιπή.

4α. Η εφαρμογή εμφανίζει μήνυμα στο χρήστη και επανέρχεται στο βήμα 2.

4. Μη λειτουργικές απαιτήσεις

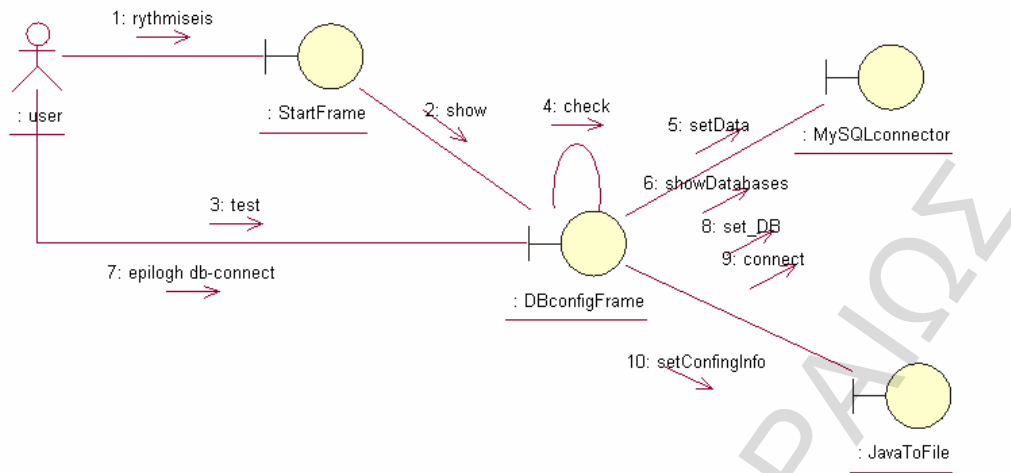
Δεν υπάρχουν για αυτή την περίπτωση χρήσης.

5. Κατάσταση εισόδου

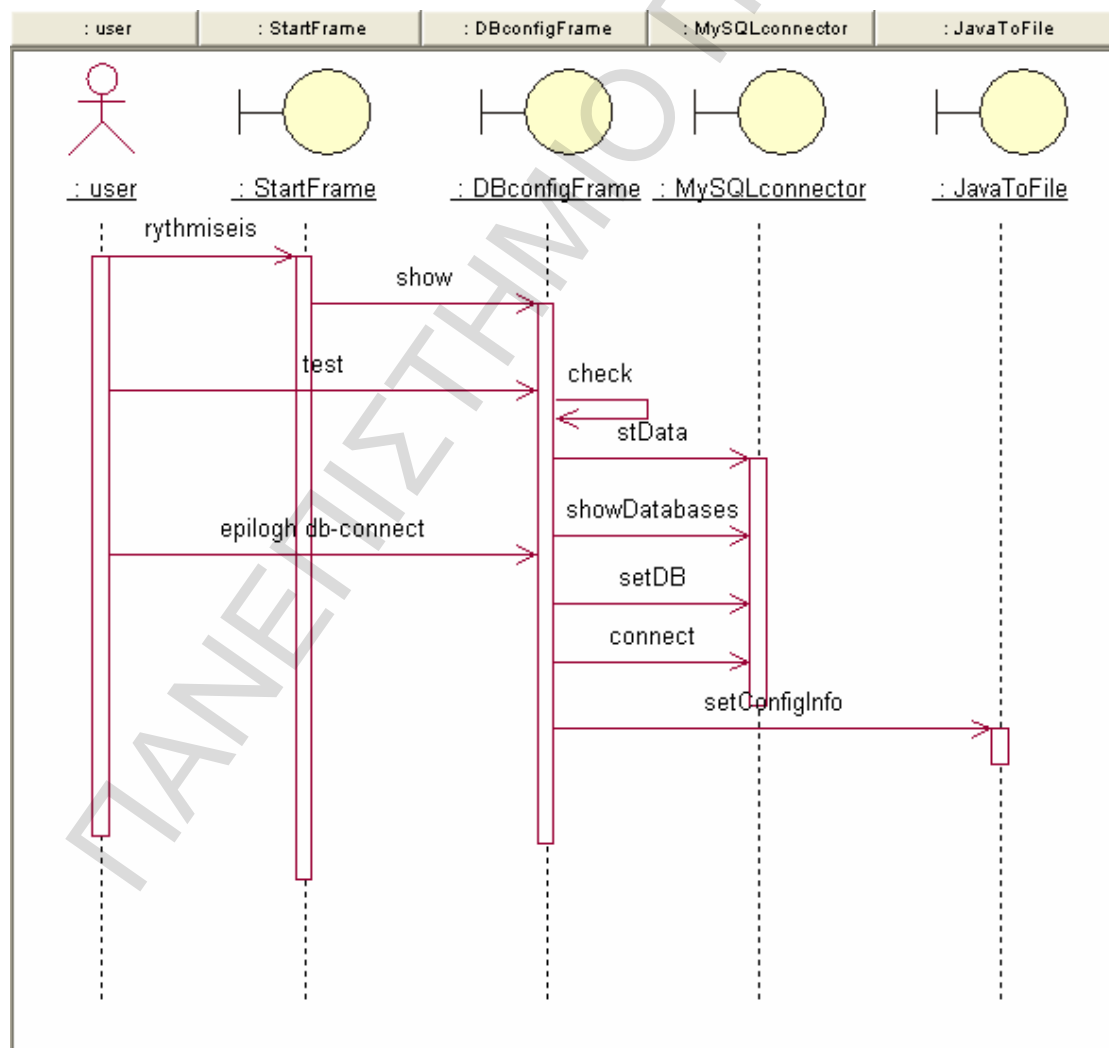
Δεν υπάρχουν ιδιαίτερες απαιτήσεις εισόδου στην περίπτωση χρήσης

6. Κατάσταση εξόδου

εμφανίζεται η ΑΡΧΙΚΗ ΟΘΟΝΗ του χρήστη.



Εικόνα 3.11 Διάγραμμα ακολουθίας της περίπτωσης χρήσης 3.2.3.5



Εικόνα 3.12 Διάγραμμα συνεργασίας της περίπτωσης χρήσης 3.2.3.5

3.2.3.6 Προδιαγραφή περίπτωσης χρήσης

1. Τίτλος περίπτωσης χρήσης

Εξαγωγή δεδομένων από την βάση δεδομένων

2. Σύντομη περιγραφή

Ο διαχειριστής εξαγάγει δεδομένα από τη βάση δεδομένων σε αρχείο λογιστικών φύλλων και σε αρχείο κειμένου

3. Ροή γεγονότων

3.1 Βασική ροή

1. Ο χρήστης από την ΑΡΧΙΚΗ ΟΘΟΝΗ επιλέγει την εργασία «ΕΙΣΟΔΟΣ ΔΙΑΧΕΙΡΙΣΤΗ»

2. Η εφαρμογή εμφανίζει την οθόνη «AdminFrame» .

3. Ο χρήστης επιλέγει την εργασία «ΑΡΧΕΙΟ EXCEL»

4. Η εφαρμογή επικοινωνεί με την βάση δεδομένων και δημιουργεί ένα αρχείο με κατάληξη .xls και όνομα την τρέχουσα ημερομηνία

5. Ο χρήστης επιλέγει την εργασία «ΑΡΧΕΙΟ ΚΕΙΜΕΝΟΥ»

6. Η εφαρμογή επικοινωνεί με την βάση δεδομένων και δημιουργεί ένα αρχείο με κατάληξη .txt και όνομα την τρέχουσα ημερομηνία

7. Ο έλεγχος επιστρέφει στο βήμα 2

3.2 Εναλλακτικές ροές

Δεν υπάρχουν για αυτή την περίπτωση χρήσης.

4. Μη λειτουργικές απαιτήσεις

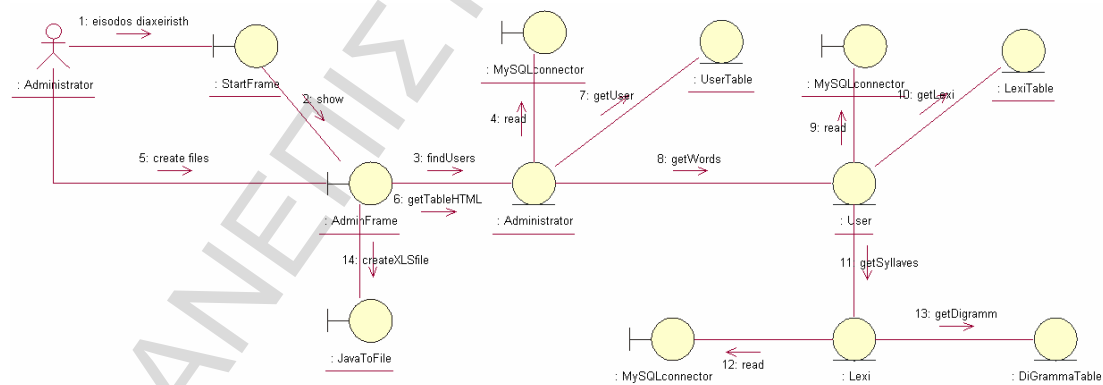
Δεν υπάρχουν για αυτή την περίπτωση χρήσης.

5. Κατάσταση εισόδου

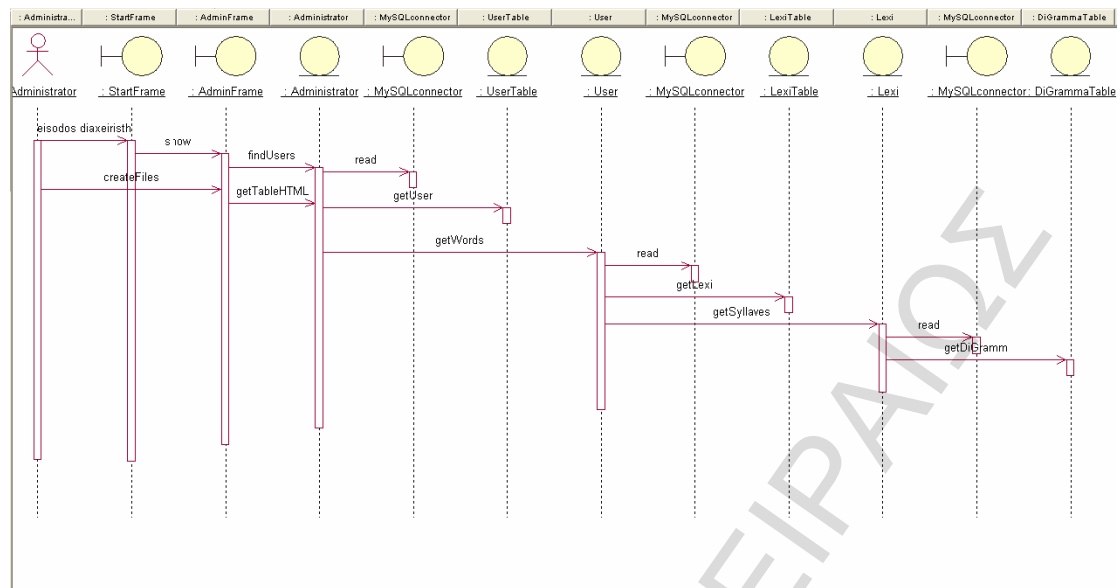
Δεν υπάρχουν ιδιαίτερες απαιτήσεις εισόδου στην περίπτωση χρήσης

6. Κατάσταση εξόδου

εμφανίζεται η ΑΡΧΙΚΗ ΟΘΟΝΗ του χρήστη.

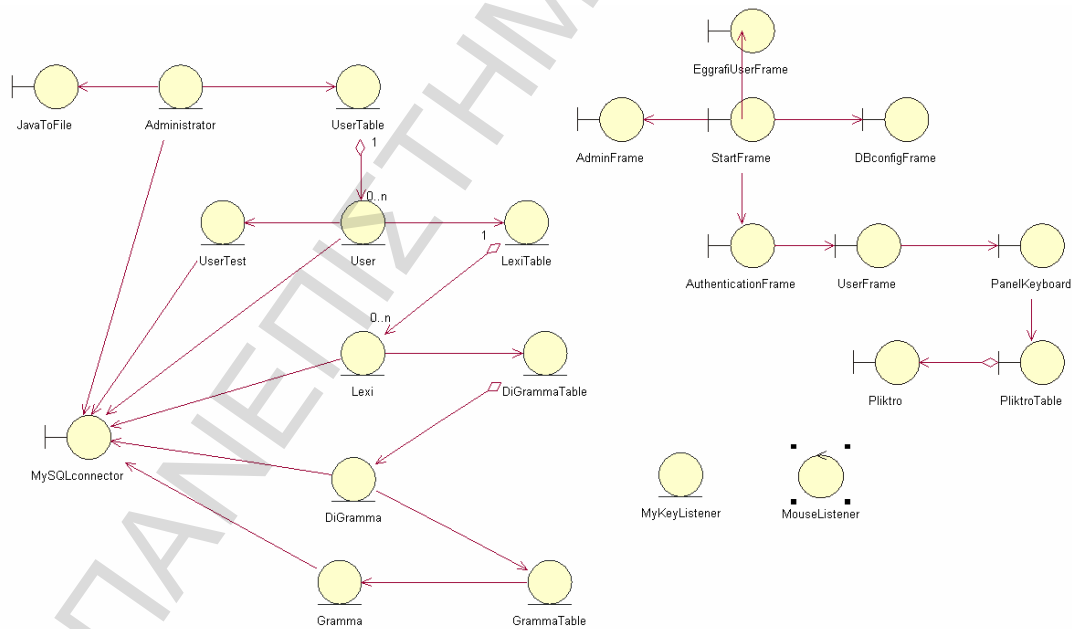


Εικόνα 3.13 Διάγραμμα ακολουθίας της περίπτωσης χρήσης 3.2.3.6



Εικόνα 3.14 Διάγραμμα συνεργασίας της περίπτωσης χρήσης 3.2.3.6

3.2.4 Διάγραμμα Κλάσεων



Εικόνα 3.15 Διάγραμμα κλάσεων της εφαρμογής

Κεφάλαιο 4

Περιγραφή προγράμματος

Το πρόγραμμα που χρησιμοποιήσαμε για την συλλογή των χαρακτηριστικών πληκτρολόγησης των χρηστών και τη δημιουργία αρχείων λογιστικών φύλλων υλοποιήθηκε με γλώσσα προγραμματισμού JAVA.

4.1 Ρυθμίσεις σύνδεσης με τη βάση δεδομένων

Κατά την πρώτη φορά εκτέλεσης του προγράμματος εμφανίζεται η οθόνη ρυθμίσεων της σύνδεσης με τη βάση δεδομένων που βρίσκεται σε έναν από τους υπολογιστές του δικτύου:

ΜΕ ΤΙ ΕΙΔΟΥΣ ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ ΘΕΛΕΤΕ ΝΑ ΣΥΝΔΕΘΕΙΤΕ; MySQL

Η ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ ΒΡΙΣΚΕΤΑΙ:

ΣΤΟΝ ΥΠΟΛΟΓΙΣΤΗ ΜΟΥ

ΣΕ ΑΛΛΟ ΥΠΟΛΟΓΙΣΤΗ ΤΟΥ ΔΙΚΤΥΟΥ

HOST IP ADDRESS: 192.168.1.64

Συνθηματικό (username): root

Κωδικός (password): *****

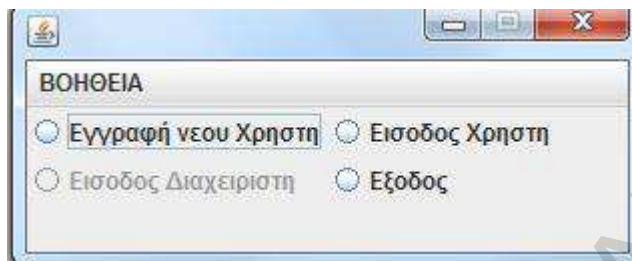
ΕΠΙΛΕΞΤΕ ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ:

TEST CONNECTION ΣΥΝΔΕΣΗ

Αφού ο χρήστης συμπληρώσει τα απαιτούμενα στοιχεία επιλέγει TEST CONNECTION για να συνδεθεί με τον MySQL SERVER. Αν η σύνδεση είναι επιτυχής

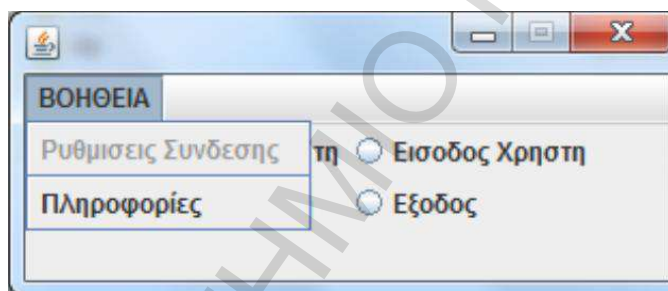
εμφανίζονται οι βάσεις δεδομένων και ο χρήστης επιλέγει την βάση δεδομένων "BIOMETRICS" και επιλέγει ΣΥΝΔΕΣΗ.

Στη συνέχεια εμφανίζεται η αρχική οθόνη του προγράμματος:



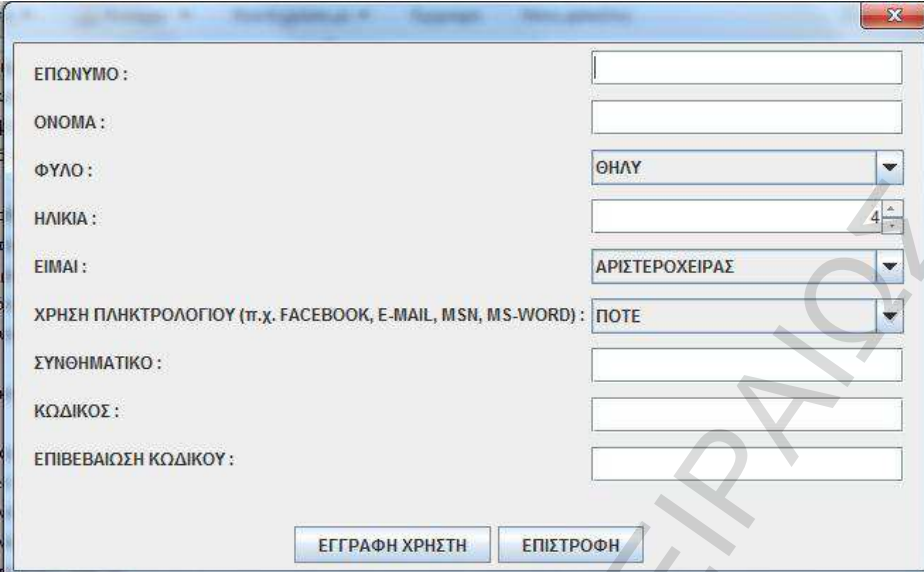
Η επιλογή «Είσοδος Διαχειριστή» είναι απενεργοποιημένη αν ο χρήστης βρίσκεται σε υπολογιστή διαφορετικό από αυτόν που φιλοξενεί τη βάση δεδομένων.

Επιλέγοντας «ΒΟΗΘΕΙΑ» από τη μπάρα μενού ο χρήστης μπορεί να ρυθμίσει τις παραμέτρους σύνδεσης με τη βάση δεδομένων όπως αναφέρθηκε προηγούμενα. Αν η επιλογή είναι απενεργοποιημένη, όπως φαίνεται στην παρακάτω εικόνα, σημαίνει ότι η σύνδεση με τη βάση δεδομένων έχει πραγματοποιηθεί.



4.2 Εγγραφή χρήστη

Όταν ο χρήστης επιλέξει την «Εγγραφή νέου χρήστη» εμφανίζεται το παράθυρο της παρακάτω εικόνας:



ΕΠΩΝΥΜΟ :

ΟΝΟΜΑ :

ΦΥΛΟ :

ΗΛΙΚΙΑ :

ΕΙΜΑΙ :

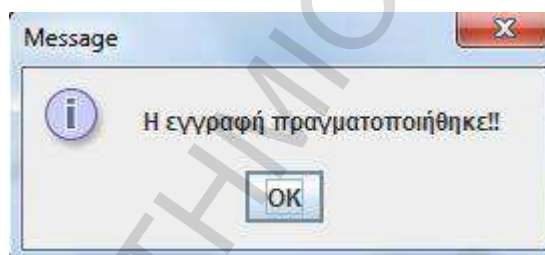
ΧΡΗΣΗ ΠΛΗΚΤΡΟΛΟΓΙΟΥ (π.χ. FACEBOOK, E-MAIL, MSN, MS-WORD) :

ΣΥΝΘΗΜΑΤΙΚΟ :

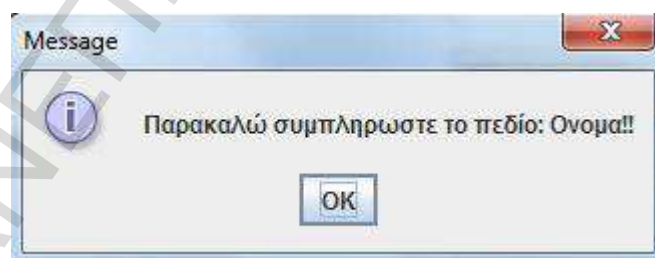
ΚΩΔΙΚΟΣ :

ΕΠΙΒΕΒΑΙΩΣΗ ΚΩΔΙΚΟΥ :

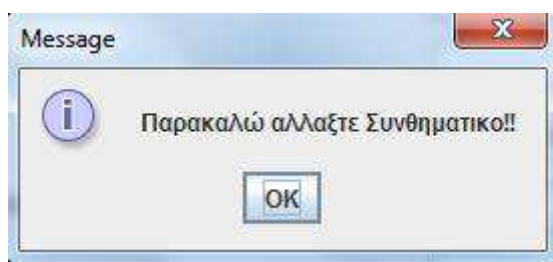
Ο χρήστης συμπληρώνει την φόρμα με ΚΕΦΑΛΑΙΟΥΣ και ΕΛΛΗΝΙΚΟΥΣ ΧΑΡΑΚΤΗΡΕΣ και επιλέγει «ΕΓΓΡΑΦΗ ΧΡΗΣΤΗ». Αν η εγγραφή πραγματοποιήθηκε επιτυχώς εμφανίζεται το παρακάτω μήνυμα:



Αν ο χρήστης παραλείψει να συμπληρώσει κάποια στοιχεία ειδοποιείται με μηνύματα σφάλματος όπως το παρακάτω:

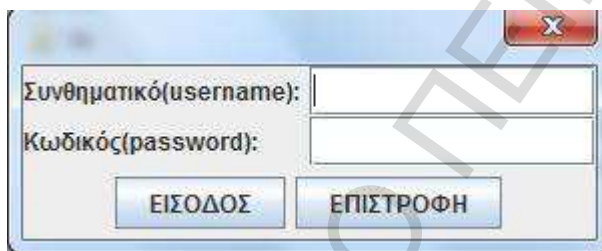


Αν ο χρήστης χρησιμοποιήσει συνθηματικό (username) που χρησιμοποιείται ήδη από άλλο χρήστη η εφαρμογή τον ενημερώνει με το παρακάτω μήνυμα :

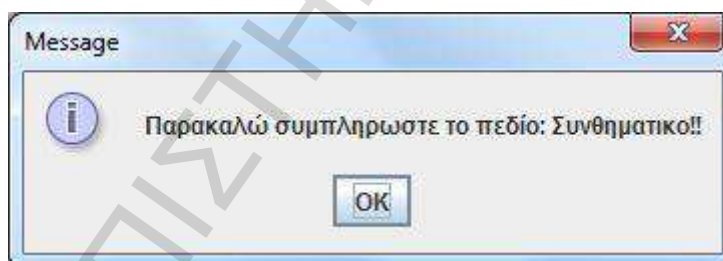


4.3 Είσοδος χρήστη

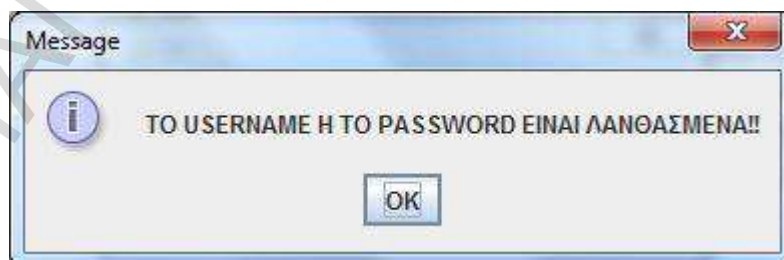
Μετά την ολοκλήρωση της εγγραφής ο χρήστης είναι έτοιμος να εισέλθει κάνοντας χρήση των συνθηματικών του επιλέγοντας από την αρχική οθόνη την επιλογή «ΕΙΣΟΔΟΣ ΧΡΗΣΤΗ» οπότε και εμφανίζεται το παρακάτω παράθυρο:



Αν παραλείψει να συμπληρώσει κάποιο από τα πεδία η εφαρμογή τον ενημερώνει με μηνύματα όπως και στην περίπτωση της εγγραφής.

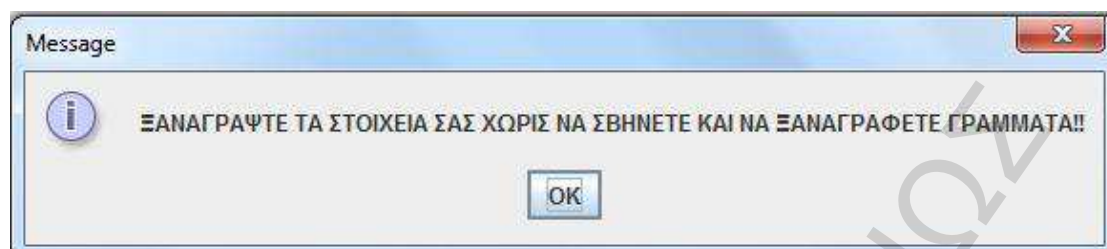


Αν κάποιο από τα συνθηματικά είναι λανθασμένο τότε εμφανίζεται το παρακάτω μήνυμα λάθους:



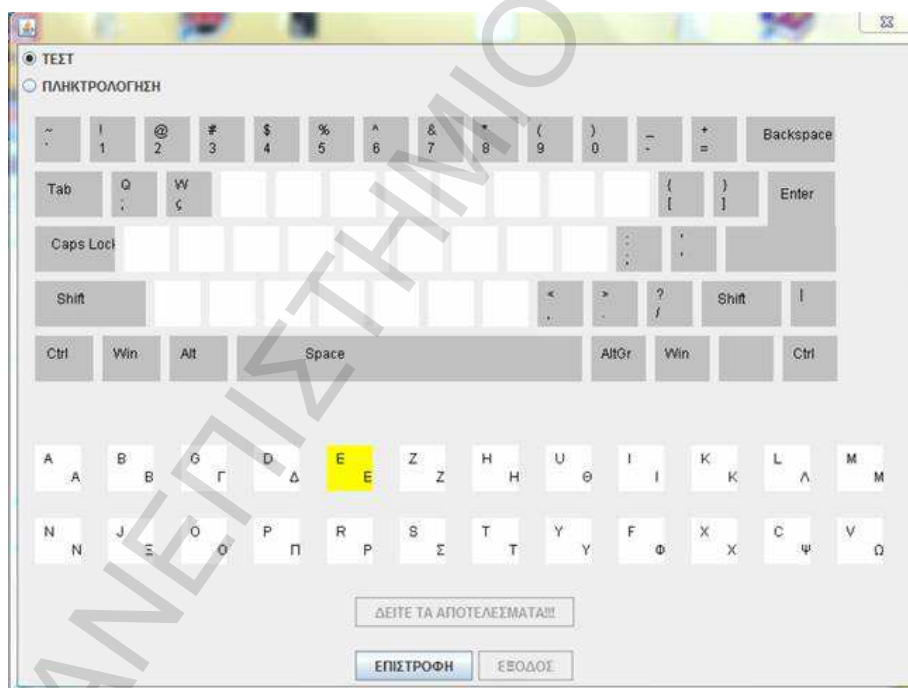
Τα συνθηματικά θα πρέπει να γραφτούν χωρίς διορθώσεις, χωρίς δηλαδή να γραφτούν και να σβηστούν γράμματα κατά την πληκτρολόγηση του συνθηματικού ή του

κωδικού. Σε περίπτωση λάθους θα πρέπει να επαναπληκτρολογηθεί το username ή το password χωρίς λάθος αλλιώς εμφανίζεται το παρακάτω μήνυμα:

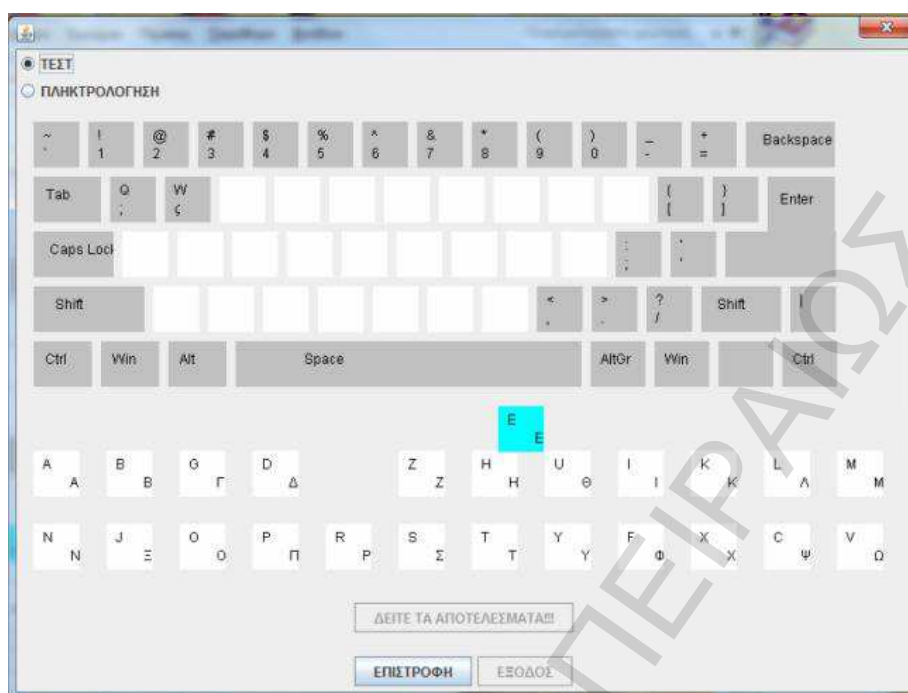


4.4 Πραγματοποίηση τεστ

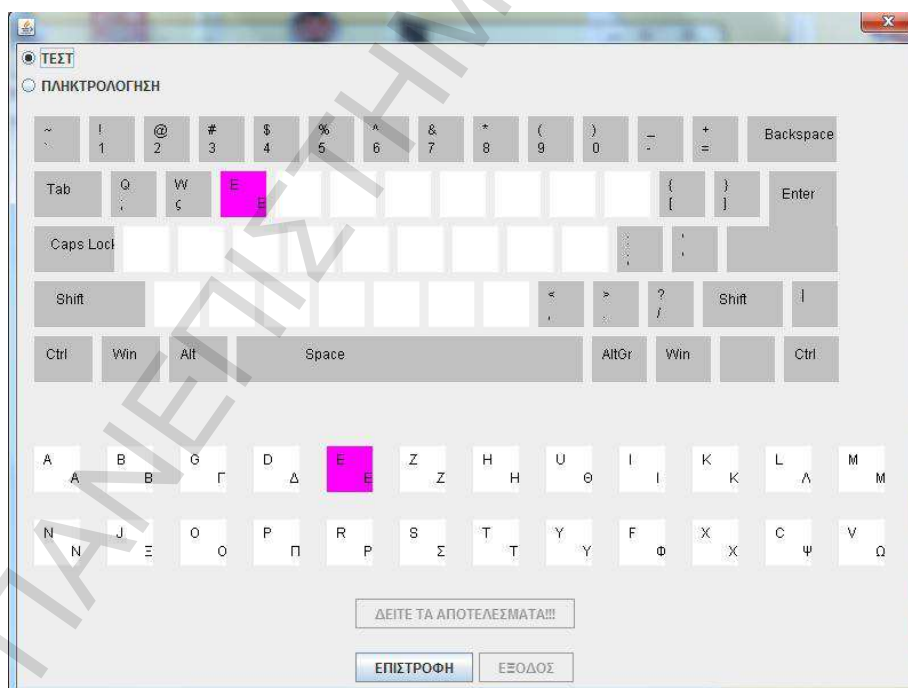
Μετά την επιτυχή επιβεβαίωση του, ο χρήστης επιλέγει το «ΤΕΣΤ» η πραγματοποίηση του οποίου είναι απαραίτητη για να ενεργοποιηθεί η διαδικασία εμφάνισης των λέξεων που πρέπει να πληκτρολογήσει. Στο παράθυρο εμφανίζεται μια απεικόνιση ενός τυπικού πληκτρολογίου και κάτω από αυτό τα γράμματα του ελληνικού αλφαβήτου. Ο χρήστης καλείται με τη χρήση ποντικιού (drag and drop) και χωρίς να συμβουλευτεί το πληκτρολόγιο να τοποθετήσει τα γράμματα στις θέσεις που πιστεύει ότι βρίσκονται στο πληκτρολόγιο:



Όταν ο κέρσορας του ποντικιού βρίσκεται πάνω από ένα γράμμα αυτό χρωματίζεται κίτρινο, γεγονός που σημαίνει ότι μπορεί να μετακινηθεί. Μόλις ο χρήστης κάνει επιλογή και σύρσιμο του γράμματος με αριστερό κλικ (drag and drop), το γράμμα χρωματίζεται γαλάζιο και μετακινείται.

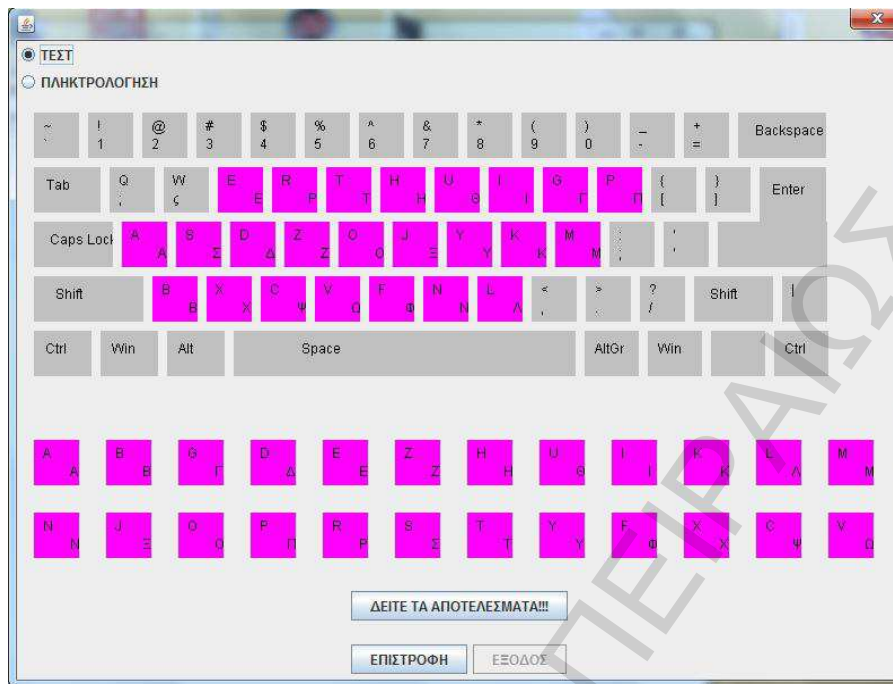


Όταν ο κέρσορας βρεθεί πάνω από κάποιο κενό (λευκό) πλήκτρο του πληκτρολόγιου, αυτό χρωματίζεται κίτρινο. Αν ο χρήστης απελευθερώσει τότε το μεταφερόμενο γράμμα, αυτό καταλαμβάνει το κενό πλήκτρο και τα δύο χρωματίζονται μωβ που σημαίνει ότι έχουν δεσμευτεί.



Αν ο χρήστης επιθυμεί να αναιρέσει την επιλογή θέσης του γράμματος θα πρέπει να κάνει διπλό αριστερό κλικ πάνω στην μωβ θέση του πληκτρολόγιου και τότε και το πλήκτρο και το γράμμα αποδεσμεύονται και μπορούν να ξαναχρησιμοποιηθούν.

Μόλις η παραπάνω διαδικασία ολοκληρωθεί και όλα τα γράμματα έχουν τοποθετηθεί σε θέσεις του πληκτρολόγιου ενεργοποιείται η επιλογή «ΔΕΙΤΕ ΤΑ ΑΠΟΤΕΛΕΣΜΑΤΑ».

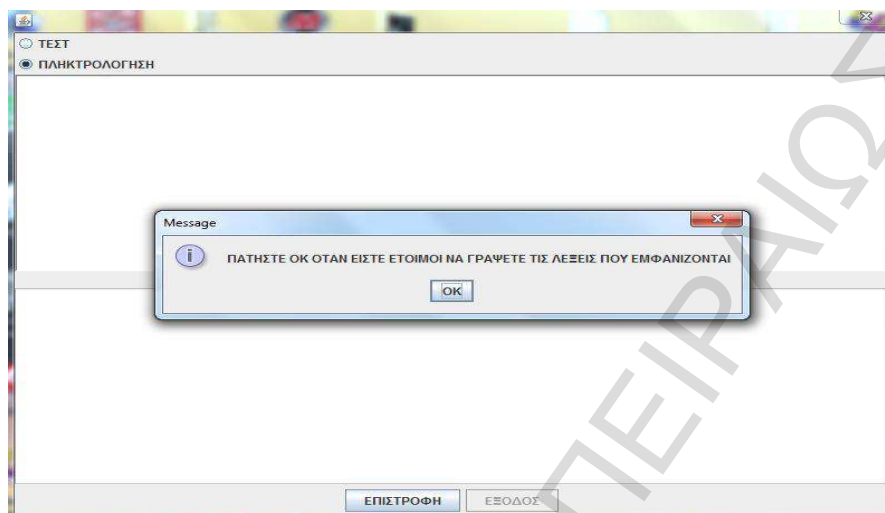


Όταν ο χρήστης επιλέξει να δει τα αποτελέσματα εμφανίζεται μήνυμα όπου φαίνεται ο χρόνος, τα λάθη και η προσπάθειες που έκανε μέχρι να ολοκληρώσει το τεστ. Εδώ υπενθυμίζεται ότι η εφαρμογή υπολογίζει όχι μόνο τα λάθη αλλά και την απόσταση των λαθών ως κριτήριο της άγνοιας του χρήστη, μιας και η τοποθέτηση ενός γράμματος σε θέση πλησίον της πραγματικής του θέσης δείχνει ότι ο χρήστης έχει γνώση που περίπου βρίσκεται το γράμμα πάνω στο πληκτρολόγιο ενώ η τοποθέτηση του σε μια απομακρυσμένη από τη σωστή θέση δείχνει ότι ο χρήστης έχει άγνοια της θέσης του γράμματος πάνω στο πληκτρολόγιο. Επίσης οι σωστές επιλογές του χρήστη εμφανίζονται πράσινες ενώ οι λανθασμένες τοποθετήσεις γραμμάτων κόκκινες.



4.5 Πληκτρολόγηση λέξεων

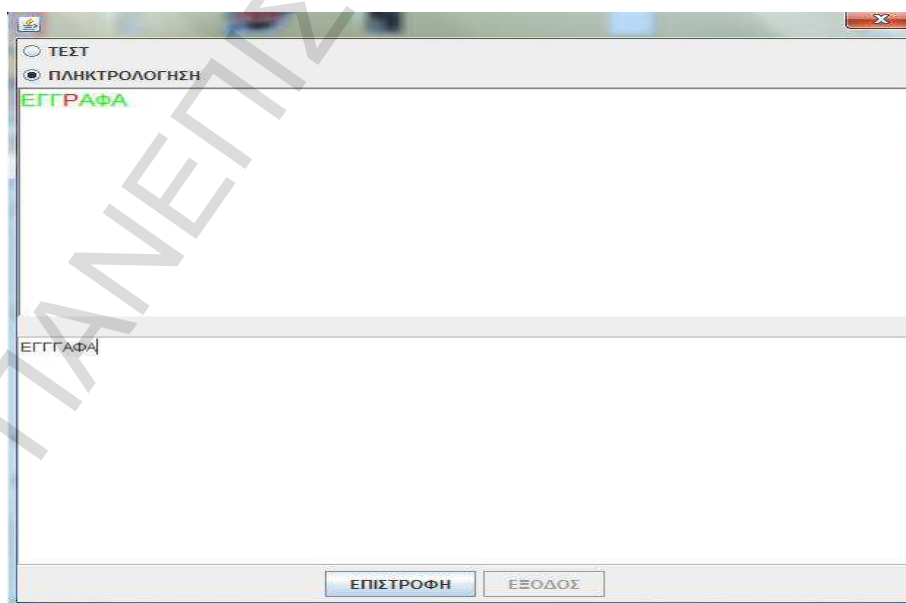
Μετά την ολοκλήρωση του τεστ ο χρήστης επιλέγει «ΠΛΗΚΤΡΟΛΟΓΗΣΗ» και εμφανίζεται το παρακάτω παράθυρο της εφαρμογής:



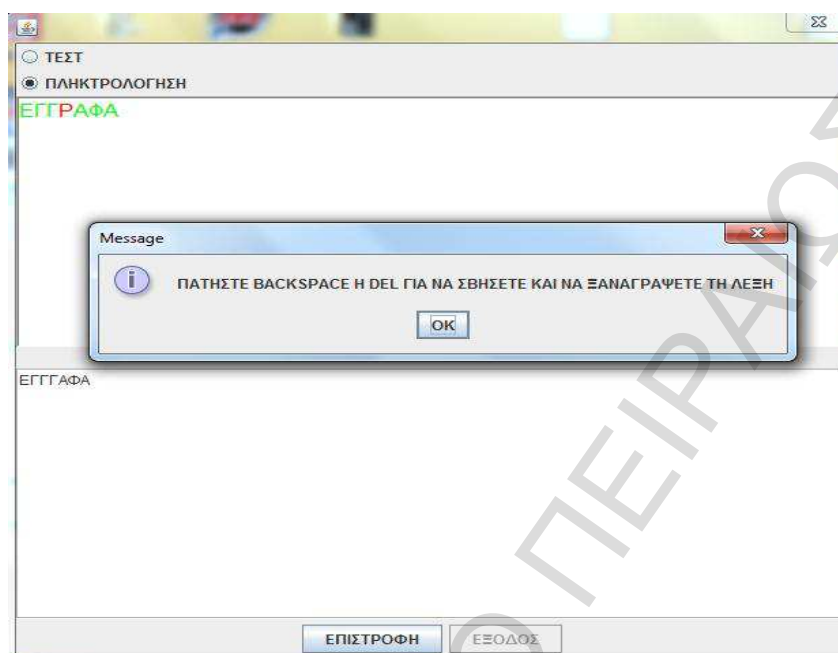
Όταν ο χρήστης επιβεβαιώσει την ετοιμότητα του στο πάνω παράθυρο εμφανίζονται με προκαθορισμένη σειρά και μια-μια οι 52 λέξεις που καλείται να πληκτρολογήσει στο κάτω παράθυρο της εφαρμογής.

Μόλις ολοκληρώσει την πληκτρολόγηση μιας λέξης χωρίς λάθος (γεγονός που επιβεβαιώνεται με τον πράσινο χρωματισμό όλων των γραμμάτων της λέξης που παρουσιάζεται στο πάνω παράθυρο), με την πίεση του ENTER ή του SPACE, τα χαρακτηριστικά της πληκτρολόγησης καταχωρούνται και εμφανίζεται η νέα λέξη στο επάνω παράθυρο και ο χρήστης επαναλαμβάνει τη διαδικασία.

Σε περίπτωση λάθους ο χρήστης ειδοποιείται ηχητικά και οπτικά (το λάθος γράμμα χρωματίζεται κόκκινο) ώστε να διακόψει την προσπάθεια, όπως φαίνεται στην παρακάτω εικόνα:



Αν προσπαθήσει να την καταχωρήσει τότε εμφανίζεται μήνυμα λάθους που τον καθοδηγεί για την επόμενη ενέργεια του.



Μόλις πραγματοποιηθεί η ορθή πληκτρολόγηση και των 52 λέξεων το παράθυρο στο οποίο πληκτρολογούσε ο χρήστης δεν είναι πια προσβάσιμο και ενεργοποιείται η επιλογή «ΕΞΟΔΟΣ» με τη χρήση της οποίας τερματίζεται η εφαρμογή.

4.6 Είσοδος διαχειριστή

Όταν η εφαρμογή είναι εγκατεστημένη στον υπολογιστή όπου βρίσκεται η βάση δεδομένων «BIOMETRICS» η επιλογή αυτή οδηγεί σε παράθυρο που επιτρέπει στον διαχειριστή του προγράμματος να δημιουργήσει στον υποκατάλογο «ΑΡΧΕΙΑ» του καταλόγου του προγράμματος αρχεία EXCEL ή/και αρχεία κειμένου με όνομα την ημερομηνία που πραγματοποιήθηκε αυτή η ενέργεια. Τα αρχεία αυτά περιέχουν τα χαρακτηριστικά των διγραμμάτων και των γραμμάτων που πληκτρολογήθηκαν από κάθε χρήστη.

Κεφάλαιο 5

Περιγραφή έρευνας

Γενικά

Ο τρόπος πληκτρολόγησης για κάθε άτομο αναλύεται στα παρακάτω χαρακτηριστικά:

- Χρονική Διάρκεια ή «χρόνος πίεσης»
Είναι το χρονικό διάστημα που μεσολαβεί από τη στιγμή έναρξης πίεσης ενός πλήκτρου ώστε να γραφτεί το αντίστοιχο γράμμα (ή σύμβολο) μέχρι τη στιγμή της απελευθέρωσης του. Οι μέχρι τώρα έρευνες δείχνουν ότι είναι το πιο σταθερό χρονικά χαρακτηριστικό του τρόπου πληκτρολόγησης
- Χρονική Καθυστέρηση ή «χρόνος πτήσης»
Είναι το χρονικό διάστημα μεταξύ διαδοχικών πληκτρολογήσεων. Ο υπολογισμός αυτού του χαρακτηριστικού μπορεί να γίνει με δύο τρόπους:
 - α) ο χρόνος πτήσης είναι το χρονικό διάστημα από τη στιγμή της απελευθέρωσης του πρώτου πλήκτρου μέχρι τη στιγμή της έναρξης πίεσης του δεύτερου.
 - β) ο χρόνος πτήσης είναι το χρονικό διάστημα από τη στιγμή έναρξης πίεσης του πρώτου πλήκτρου ως τη στιγμή έναρξης της πίεσης του δεύτερου.
 Η σχέση που συνδέει τους χρόνους πτήσης των περιπτώσεων (α) και (β) είναι:

$$t_{fly_b} = t_{press_1} + t_{fly_a}$$

όπου:

t_{fly_b} : ο χρόνος πτήσης όπως περιγράφεται στην περίπτωση (β)

t_{fly_a} : ο χρόνος πτήσης όπως περιγράφεται στην περίπτωση (α)

t_{press_1} : ο χρόνος πίεσης του 1^{ου} γράμματος του διγράμματος

Στην περίπτωση της σειριακής πληκτρολόγησης, όπως πχ συμβαίνει όταν ο χρήστης πληκτρολογεί με ένα δάχτυλο, θα πρέπει να χρησιμοποιείται ο πρώτος υπολογισμός, γιατί αν χρησιμοποιηθεί ο δεύτερος τότε ο χρόνος πτήσης περιέχει και τον χρόνο πίεσης του πρώτου γράμματος.

Αντίθετα αν ο χρήστης πληκτρολογεί με περισσότερα από 2 δάχτυλα, είναι φρόνιμο να χρησιμοποιείται ο δεύτερος τρόπος υπολογισμού, καθώς αυτός περιγράφει καλύτερα τη χρονική ακολουθία των γεγονότων και δεν επηρεάζεται από το χρόνο πίεσης του πρώτου πλήκτρου.

Όπως προείπαμε ο τρόπος πληκτρολόγησης είναι ένα ψυχολογικό βιομετρικό χαρακτηριστικό. Κατά τη διάρκεια πληκτρολόγησης μιας λέξης ή μιας πρότασης όπου ο χρήστης βρίσκεται κάτω από την ίδια ψυχολογική και συναισθηματική κατάσταση και οι συνθήκες του περιβάλλοντος παραμένουν σταθερές, από ποιους παράγοντες εξαρτάται ο χρόνος πτήσης μεταξύ των διαφορετικών διγράμμάτων που πληκτρολογούνται; Που οφείλεται η διακύμανση των τιμών που παρατηρούνται στους χρόνους πτήσης του ίδιου διγράμματος;

Γενικά θα μπορούσαμε να θεωρήσουμε ότι ο χρόνος πτήσης μεταξύ δύο διαδοχικών πίεσεων πλήκτρων εξαρτάται από τον χρόνο αναζήτησης του δεύτερου γράμματος και από το χρόνο κίνησης του δακτύλου προς το δεύτερο γράμμα. Έτσι:

$$t_{fly} = f(t_{αναζ}, t_{move})$$

όπου:

$t_{αναζ}$: ο χρόνος αναζήτησης είναι το χρονικό διάστημα που ο χρήστης σκέφτεται-προσπαθεί να εντοπίσει τη θέση του επόμενου γράμματος

t_{move} : ο χρόνος κίνησης είναι το χρονικό διάστημα που το δάκτυλο κινείται προς το επόμενο γράμμα.

Ο χρόνος αναζήτησης εξαρτάται από την εμπειρία του χρήστη η οποία είναι προϊόν προηγούμενων αναζητήσεων του ίδιου γράμματος και εμπεριέχει όλους εκείνους τους παράγοντες που επηρεάζουν το βαθμό συγκέντρωσης του. Κατά συνέπεια ο χρόνος αναζήτησης, εφ' όσον ο χρήστης έχει πληκτρολογήσει ικανοποιητικό αριθμό λέξεων, πιθανόν να εξαρτάται από τη συχνότητα εμφάνισης του δεύτερου γράμματος στη γλώσσα.

Ο χρόνος κίνησης αναφέρεται στη μηχανική κίνηση των δακτύλων και εξαρτάται από τον τρόπο γραφής και την σχετική θέση των δύο γραμμάτων στο πληκτρολόγιο.

Οι δύο αυτοί χρόνοι δεν παρουσιάζουν σειριακή ακολουθία, καθώς ο χρήστης μπορεί, ενώ αναζητά το επόμενο γράμμα, το χέρι του να κινείται προς αυτή την κατεύθυνση. Συνεπώς, ο χρόνος πτήσης δεν μπορεί να υπολογιστεί από το αλγεβρικό άθροισμα τους.

5.1 Προεργασία

Ενώ τα στατιστικά της ελληνικής γλώσσας όσον αφορά τα γράμματα είναι γνωστά, η παρουσία των διγραμμάτων δεν είναι σταθερή και επηρεάζεται από πολλούς παράγοντες. Από διάφορα κείμενα μετρήθηκαν 208.752 διγράμματα και υπολογίστηκαν τα ποσοστά εμφάνισης για κάθε ένα από τα 576 διγράμματα που προκύπτουν από το συνδυασμό των 24 γραμμάτων. Τα διγράμματα ομαδοποιήθηκαν ανάλογα με τη θέση των γραμμάτων στο πληκτρολόγιο. Στη συνέχεια επιλέχθηκαν τα διγράμματα με τη μεγαλύτερη συχνότητα εμφάνισης, τα αντίθετα τους διγράμματα, διγράμματα που εμφανίζουν παρόμοια χαρακτηριστικά με τα συχνότερα διγράμματα και διγράμματα που αποτελούνται από ίδια γράμματα. Τα διγράμματα που επιλέχθηκαν ήταν 63, 43 εκ των οποίων συλλαβές ενώ τα 20 μη συλλαβές. Το συνολικό ποσοστό εμφάνισης των διγραμμάτων που επιλέχθηκαν προς μελέτη αποτελεί το 34,3% , δηλαδή 1 στα 3 διγράμματα που έχουν πληκτρολογήσει στον παρελθόν οι χρήστες είναι κάποιο από τα επιλεγμένα.

5.2 Οι λέξεις του τεστ

Οι λέξεις του τεστ επιλέχθηκαν ώστε να περιέχονται σε αυτές όσο το δυνατόν συχνότερα τα 63 διγράμματα που θα μελετηθούν στατιστικά. Από τα διγράμματα αυτά τα 43 είναι συλλαβές και τα 20 είναι μη συλλαβές.

Η πληκτρολόγηση έγινε με κεφαλαία γράμματα για να αποφευχθεί η χρήση τόνων και να αντικατασταθούν τα σ,ς με το Σ. έτσι χρησιμοποιήθηκαν μόνο τα πλήκτρα των 24 γραμμάτων του ελληνικού αλφαβήτου.

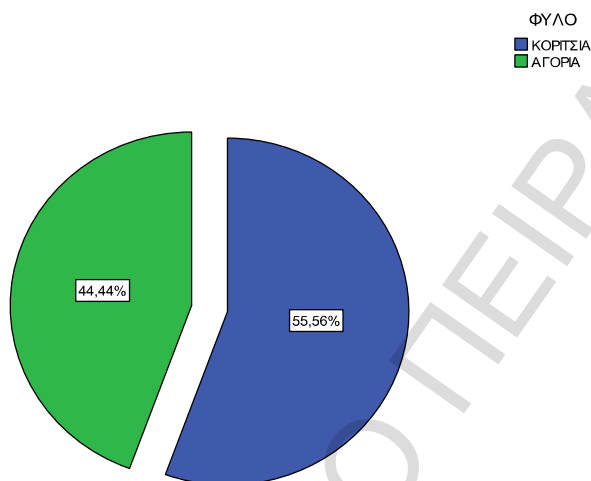
5.3 Διεξαγωγή έρευνας

Για να μελετηθεί ο τρόπος γραφής των χρηστών η εφαρμογή χρησιμοποιήθηκε από 9 χρήστες, μαθητές του τμήματος Πληροφορικής του 1^{ου} ΕΠΑΛ Καισαριανής. Αυτοί πληκτρολόγησαν τις 52 λέξεις του τεστ με το δικό τους τρόπο πληκτρολόγησης, 3 φορές. Τα χρονικά διαστήματα μεταξύ των τεστ ήταν μία εβδομάδα ανάμεσα στο 1^ο και το 2^ο και 2 εβδομάδες ανάμεσα στο 2^ο και το 3^ο.

5.4 Εθελοντές

Οι μαθητές που συμμετείχαν στα τεστ ήταν φυσικοί ομιλητές της ελληνικής γλώσσας και είχαν γενικά ετερογενείς δεξιότητες πληκτρολόγησης. Οι 4 από τους 9 ήταν αγόρια και οι 5 ήταν κορίτσια, ενώ ένα αγόρι και 1 κορίτσι ήταν αριστερόχειρες. Το κάθε τεστ πραγματοποιήθηκε στη διάρκεια μιας διδακτικής ώρας στο εργαστήριο πληροφορικής και κάθε μαθητής επανέλαβε το τεστ χρησιμοποιώντας κάθε φορά τον ίδιο υπολογιστή.

Οι χρήστες δεν γνώριζαν τον σκοπό του πειράματος ώστε να μην επηρεαστεί ο τρόπος και ο ρυθμός πληκτρολόγησης τους.



Εικόνα 5.1 Γραφική αναπαράσταση του ποσοστού του φύλλου των χρηστών

5.5 Συλλογή δεδομένων

Κατά την πληκτρολόγηση των λέξεων η κάθε λέξη του πειράματος απεικονιζόταν στην οθόνη της εφαρμογής και από κάτω φαινόταν η αναπαραγωγή της λέξης από το χρήστη, γράμμα προς γράμμα, καθώς αυτός την πληκτρολογούσε. Μετά την ορθή καταχώρηση κάθε λέξης από το χρήστη η εφαρμογή παρουσίαζε την επόμενη.

Λανθασμένες λέξεις δεν ήταν αποδεκτές. Σε περίπτωση λάθους η εφαρμογή ειδοποιούσε το χρήστη οπτικά και ακουστικά μέχρι ο χρήστης να διαγράψει πλήρως τη λέξη και να επαναλάβει την πληκτρολόγησή της.

Οι χρήστες ήταν ελεύθεροι να επιλέξουν τον ρυθμό πληκτρολόγησης κάθε λέξης ή και να σταματήσουν όποτε αυτοί επιθυμούσαν ακόμα και κατά τη διάρκεια πληκτρολόγησης.

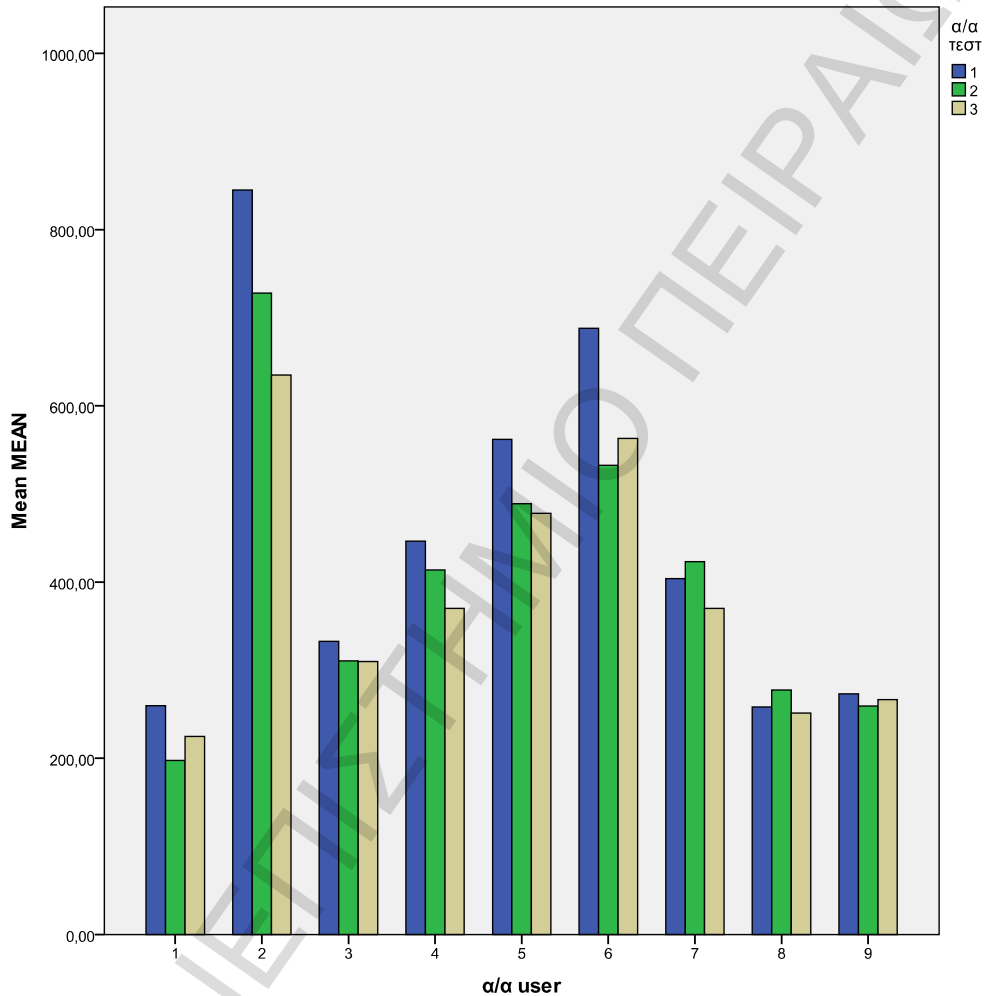
5.6 Στατιστική επεξεργασία δεδομένων

Μετά την ολοκλήρωση των πειραμάτων η εφαρμογή δημιουργεί αρχεία λογιστικών φύλλων ή αρχεία κειμένου ώστε να χρησιμοποιηθούν για την εισαγωγή δεδομένων στο SPSS 17 με το οποίο έγινε η στατιστική επεξεργασία τους. Οι έλεγχοι που ακολουθούν πραγματοποιήθηκαν με επίπεδο σημαντικότητας 5%.

Κεφάλαιο 6

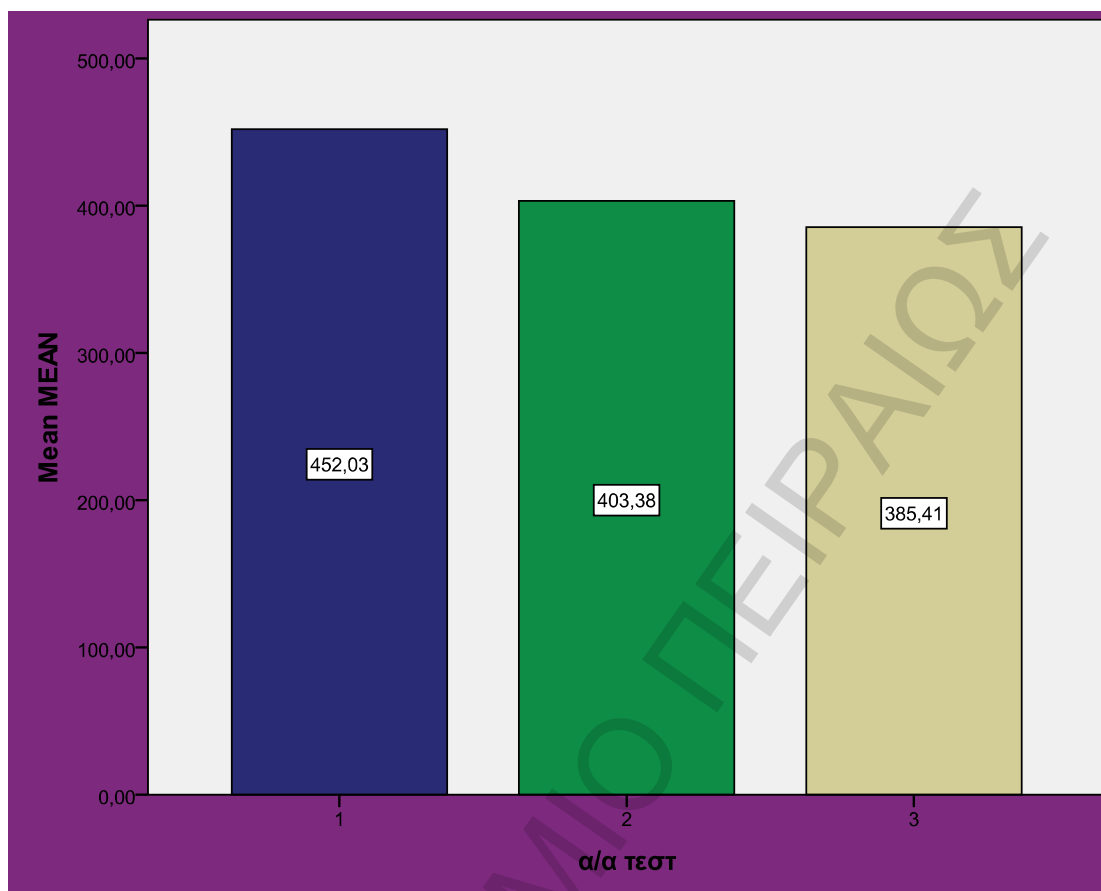
6.1 Στατιστικοί έλεγχοι και αποτελέσματα

Στο παρακάτω διάγραμμα παρουσιάζονται οι μέσοι χρόνοι πτήσης που σημείωσαν οι χρήστες σε κάθε μια από τις 3 δοκιμασίες. Όπως φαίνεται, οι χρήστες έχουν διαφορετικές δεξιότητες πληκτρολόγησης και οι πιο άπειροι χρήστες εμφανίζουν σημαντική διαφοροποίηση στους μέσους χρόνους πτήσης στις 3 επαναλήψεις της διαδικασίας.



Εικόνα 6.1 Γράφημα του μ.ο. του χρόνου πτήσης σε κάθε ένα τεστ για κάθε χρήστη

Στο σύνολο των χρηστών παρατηρείται μικρή μείωση του μέσου χρόνου πτήσης μεταξύ των 3 τεστ όπως φαίνεται στο παρακάτω διάγραμμα



Εικόνα 6.2 Γράφημα του μ.ο. του χρόνου πτήσης όλων των χρηστών σε κάθε ένα τεστ

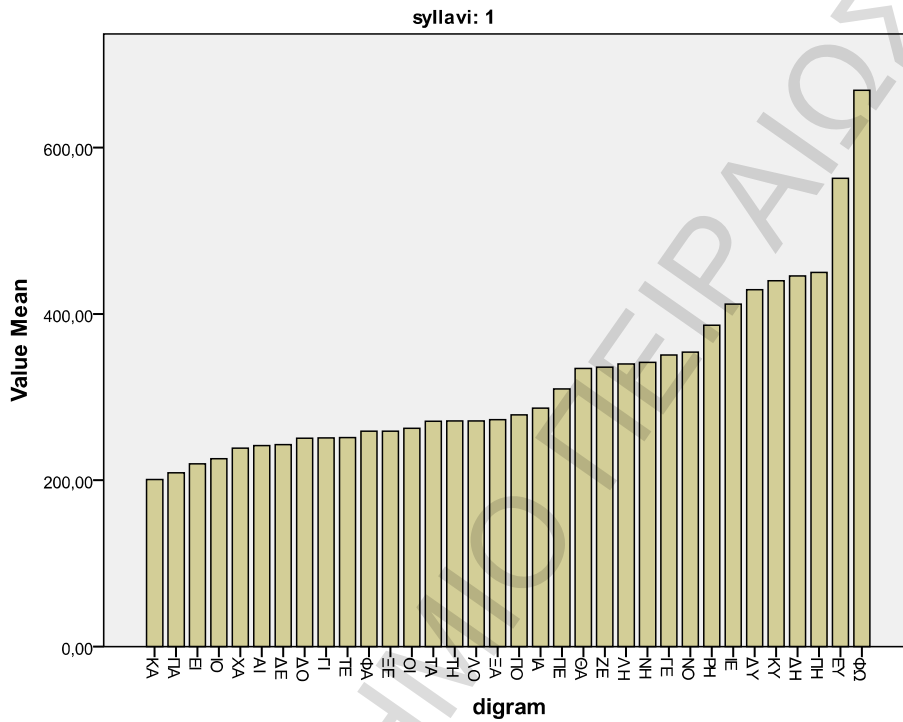
Ο μέσος χρόνος πτήσης των διγραμμάτων των χρηστών ταξινομημένος κατά αύξουσα σειρά φαίνεται στον παρακάτω πίνακα:

digram	N	Mean	Std. Deviation	Minimum	25	50	75	Maximum
ΚΑ	187	200,84	132,970	70	124,00	156,00	234,00	889
ΠΑ	80	209,01	130,277	62	129,25	163,50	248,00	785
ΕΙ	215	219,76	146,877	63	142,00	175,00	243,00	1046
ΙΟ	83	225,86	127,674	123	156,00	187,00	234,00	812
ΧΑ	109	238,62	225,957	79	134,50	203,00	271,00	2246
ΑΙ	105	241,76	288,242	62	110,00	156,00	214,50	1706
ΔΕ	81	242,86	109,009	109	172,00	210,00	265,50	718
ΔΟ	107	250,55	193,849	78	150,00	187,00	250,00	1269
ΓΙ	109	251,08	211,446	93	159,50	210,00	265,00	2002
ΤΕ	81	251,42	203,099	67	125,00	188,00	281,00	1305
ΦΑ	82	259,13	218,489	78	125,75	179,50	301,50	1290
ΞΕ	110	259,14	251,905	63	124,00	164,00	272,75	1530

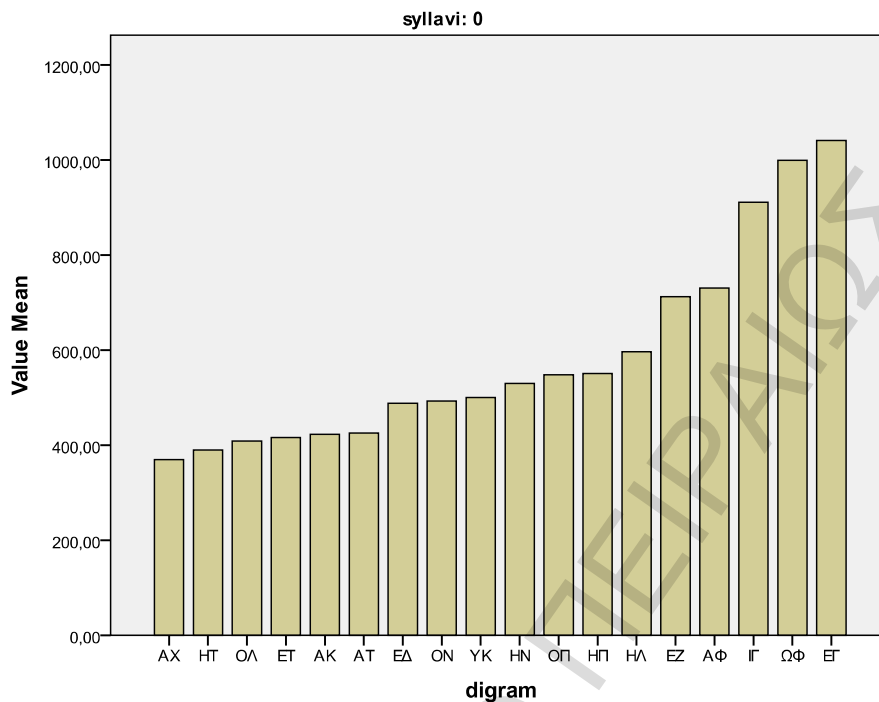
ΟΙ	163	262,47	261,156	47	156,00	187,00	234,00	2013
ΤΑ	162	271,04	233,016	62	125,00	171,00	326,50	1654
ΤΗ	137	271,17	233,880	109	165,00	200,00	263,00	1668
ΛΟ	108	271,24	204,959	124	170,00	187,00	257,50	1262
ΞΑ	53	272,98	293,274	71	125,00	156,00	272,00	1398
ΠΟ	110	278,63	368,550	33	125,75	160,00	224,25	2623
ΙΑ	80	286,75	227,774	78	140,25	185,50	359,00	1045
ΠΕ	135	309,66	314,880	78	140,00	172,00	343,00	1736
ΘΑ	54	334,30	374,994	93	140,75	187,00	359,00	1778
ΖΕ	81	335,89	115,097	110	265,00	312,00	378,00	817
ΛΗ	55	339,76	260,790	94	265,00	312,00	378,00	1670
ΝΗ	161	341,75	316,026	62	170,00	202,00	333,00	1707
ΓΕ	107	350,61	330,314	57	140,00	218,00	422,00	1857
ΝΟ	216	353,91	273,837	31	206,50	275,00	387,00	2450
ΑΧ	55	370,20	274,968	80	218,00	297,00	391,00	1345
ΡΗ	135	386,45	315,551	62	196,00	280,00	466,00	1811
ΗΤ	81	389,86	292,023	78	196,00	280,00	466,00	1449
ΟΛ	111	408,85	342,068	130	187,00	250,00	498,00	1736
ΙΕ	81	411,74	295,971	100	203,00	312,00	493,50	1303
ΕΤ	81	416,56	296,396	120	200,00	332,00	453,00	1248
ΑΚ	108	423,35	316,478	63	200,75	312,50	606,75	1717
ΑΤ	109	426,07	328,481	62	210,50	297,00	546,00	1872
ΔΥ	53	429,25	383,495	62	210,50	297,00	546,00	1872
ΚΥ	78	440,01	395,861	109	210,50	297,00	546,00	3340
ΔΗ	54	445,81	353,806	78	181,75	293,00	644,00	1388
ΠΗ	81	449,88	298,578	109	265,00	359,00	553,50	1681
ΕΔ	54	488,41	447,978	187	234,00	320,00	482,75	2153
ΟΝ	136	493,23	339,320	187	234,00	332,00	698,00	2090
ΥΚ	55	500,31	432,113	132	279,00	375,00	538,00	2729
ΗΝ	81	529,99	436,097	130	187,00	343,00	870,50	1610
ΟΠ	55	548,60	563,314	93	200,00	248,00	718,00	2634
ΗΠ	54	551,33	256,758	203	359,00	495,50	691,75	1185
ΕΥ	82	563,28	496,367	90	238,50	400,00	686,50	2583
ΗΛ	81	596,88	403,556	125	359,00	468,00	696,00	2403
ΦΩ	54	669,07	702,334	125	204,75	481,00	835,25	4071
ΕΖ	54	712,76	366,542	150	421,00	631,00	913,00	1900
ΑΦ	82	730,93	623,490	78	296,75	569,00	979,75	3338
ΙΓ	53	911,17	655,949	124	421,50	748,00	1255,50	2948

ΩΦ	55	999,40	603,552	160	406,00	890,00	1513,00	2418
ΕΓ	56	1041,07	1616,710	94	298,50	539,50	957,75	9189

Ο παραπάνω πίνακας απεικονίζεται γραφικά στα παρακάτω 2 γραφήματα για τα διγράμματα που είναι συλλαβές της ελληνικής γλώσσας και για αυτά που δεν είναι.



Εικόνα 6.3 Γράφημα του μ.ο. του χρόνου πτήσης όλων των χρηστών για τα διγράμματα που είναι συλλαβές



Εικόνα 6.4 Γράφημα του μ.ο. του χρόνου πτήσης όλων των χρηστών για τα διγράμματα που δεν είναι συλλαβές

Με μια πρώτη ματιά, ο μέσος χρόνος πτήσης των συλλαβών είναι μικρότερος από το μέσο χρόνο πτήσης των μη συλλαβών. Ο συντελεστής γραμμικής συσχέτισης μεταξύ του μέσου χρόνου πτήσης και του ποσοστού εμφάνισης του 2^{ου} γράμματος του στη γλώσσα έδειξε υψηλή αρνητική συσχέτιση.

$$r = -0.83$$

Το γεγονός σημαίνει ότι όσο μεγαλύτερο είναι το ποσοστό εμφάνισης του δεύτερου γράμματος τόσο μικραίνει ο μέσος χρόνος πτήσης του διγράμματος και μάλιστα από την γραμμική παλινδρόμηση προκύπτει ότι προσεγγιστικά ο μέσος χρόνος πτήσης ενός διγράμματος i υπολογίζεται από την παρακάτω εξίσωση η οποία ερμηνεύει το 72,2% της μεταβλητότητας του μοντέλου.

$$\bar{t}_{fly_i} = \frac{1}{n} \sum_{j=1}^n t_{fly_{ij}} = e^{6.609 - 0.078 * x - 0.234 * y}$$

όπου :

x = ποσοστό εμφάνισης 2^{ου} γράμματος και

y = αν είναι συλλαβή ή όχι

Αξίζει εδώ να σημειωθεί ότι η μέση τιμή του χρόνου πτήσης κάθε διγράμματος δεν εξαρτάται από τον ελάχιστο χρόνο πτήσης που σημειώθηκε μεταξύ των χρηστών. Αν θεωρήσουμε ότι ο ελάχιστος χρόνος πτήσης προσεγγίζει ικανοποιητικά το χρόνο κίνησης, τότε ο μέσος χρόνος πτήσης κάθε διγράμματος εξαρτάται αποκλειστικά από το χρόνο αναζήτησής του.

Επίσης η πολλαπλή γραμμική παλινδρόμηση της μέσης τιμής του χρόνου πτήσης των διγραμμάτων για κάθε χρήστη ξεχωριστά έδειξε ότι ο χρόνος πτήσης κάθε διγράμματος μειώνεται όσο αυξάνεται η συχνότητα εμφάνισης του 2^{ου} γράμματος του στα στατιστικά της γλώσσας και απ το αν πρόκειται για συλλαβή. Επίσης εξαρτάται από την ελάχιστη τιμή του χρόνου πτήσης που έχει επιτύχει κάθε χρήστης κατά την πληκτρολόγηση του συγκεκριμένου

διγράμματος. Η μέση τιμή του χρόνου πτήσης κάθε διγράμματος i ενός χρήστη j ακολουθούσε τον παρακάτω γενικό τύπο:

$$\bar{t}_{fly_{ij}} = e^{a-b_1x-b_2y} (t_{min_{ij}})^k$$

όπου :

x = ποσοστό εμφάνισης 2^{ου} γράμματος

y = αν είναι συλλαβή ή όχι και

$k \in \mathfrak{R}$

Οι σταθερές a, b_1, b_2, k ήταν διαφορετικές για κάθε χρήστη αλλά για όλους τους χρήστες η παραπάνω εξίσωση γίνεται, ερμηνεύοντας το 64,5% της μεταβλητότητας του μοντέλου:

$$\bar{t}_{fly_{ij}} = e^{2.061-0.016x-0.092y} (t_{min_{ij}})^{0.769}$$

Οι Gaines et al χρησιμοποίησαν τους λογάριθμους των χρόνων πτήσης των διγραμμάτων που πληκτρολόγησαν οι χρήστες καθώς παρατήρησαν ότι οι λογάριθμοι αυτοί ακολουθούν κανονική κατανομή [6]. Στις παραπάνω εξισώσεις, χρησιμοποιήθηκε ο λογάριθμος της μέσης τιμής του χρόνου πτήσης καθώς αυτός πράγματι εμφανίζει κανονική κατανομή, γεγονός που συμφωνεί με τις προηγούμενες έρευνες.

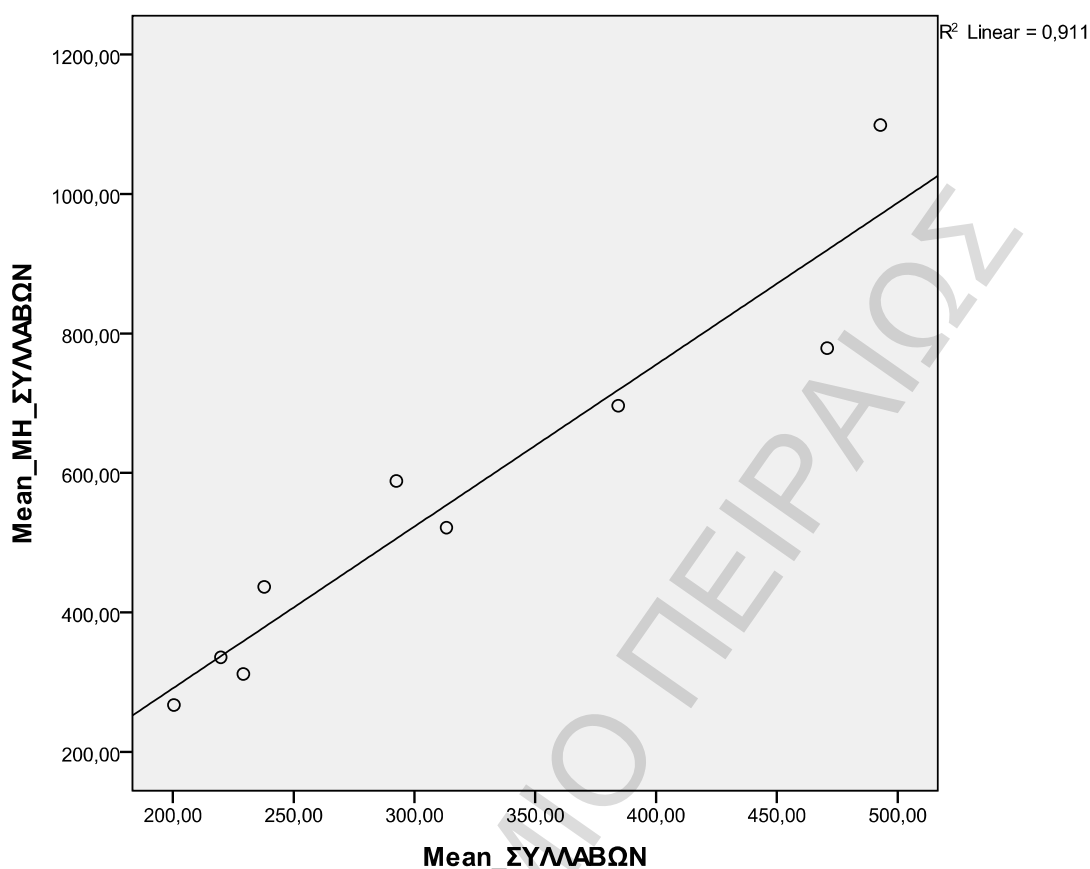
Ο μέσος χρόνος πτήσης των συλλαβών που πληκτρολόγησε συνολικά κάθε χρήστης εμφάνισε υψηλή συσχέτιση με το μέσο χρόνο πτήσης των διγραμμάτων που δεν ήταν συλλαβές, καθώς ο συντελεστής συσχέτισης βρέθηκε 0,989. Η γραμμική παλινδρόμηση έδειξε ότι ο μέσος χρόνος συλλαβών είναι:

$$y = 1.826 * x$$

όπου :

y : ο μέσος χρόνος πτήσης των μη συλλαβών κάθε χρήστη και

x : ο μέσος χρόνος πτήσης των συλλαβών κάθε χρήστη



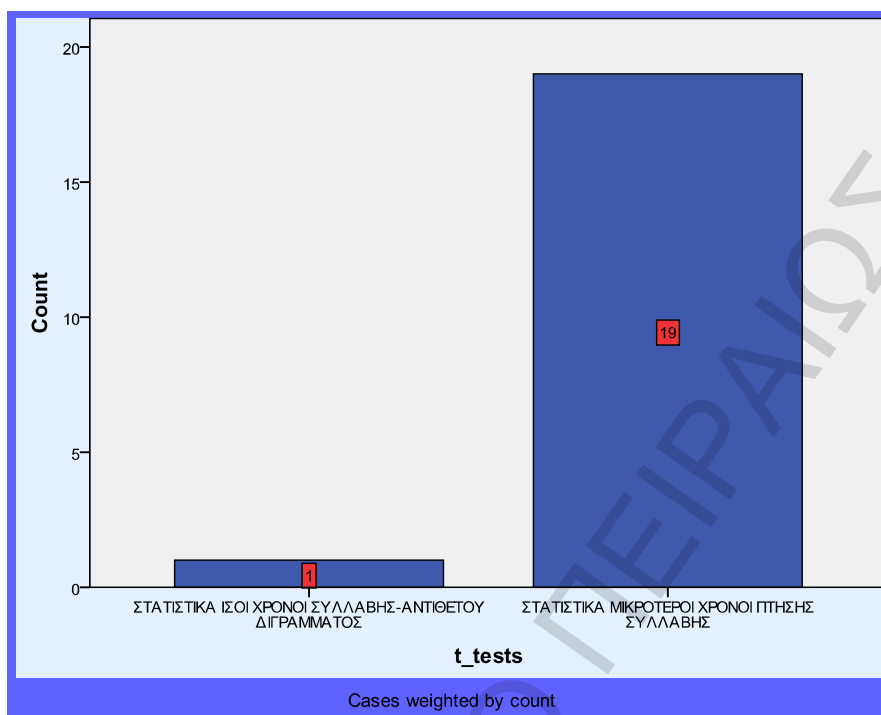
Εικόνα 6.5 Γράφημα διασποράς μεταξύ του μ.ο. του χρόνου πτήσης των συλλαβών και των μη συλλαβών για κάθε χρήστη

Η διαδικασία επαναλήφθηκε και για κάθε μια από τις 3 ημερομηνίες που πραγματοποιήθηκε το τεστ οδηγώντας στα ίδια περίπου αποτελέσματα με το συντελεστή του μέσου χρόνου πτήσης των συλλαβών να κυμαίνεται από 1,794-1,882.

Στα ίδια αποτελέσματα οδήγησε και η πραγματοποίηση t-test μεταξύ των μέσων χρόνων πτήσης των συλλαβών και των αντίθετων διγραμμάτων τους στο σύνολο των χρηστών και σε κάθε χρήστη ξεχωριστά.

Άρα ο μέσος χρόνος πτήσης των διγραμμάτων που δεν είναι συλλαβές είναι στατιστικά μεγαλύτερος από το μέσο χρόνο πτήσης των διγραμμάτων που είναι συλλαβές για το μέσο χρήστη και μάλιστα είναι 1,8 φορές μεγαλύτερος. Ο συντελεστής όμως αυτός φαίνεται να μειώνεται όσο αυξάνει η εμπειρία και η ταχύτητα του χρήστη.

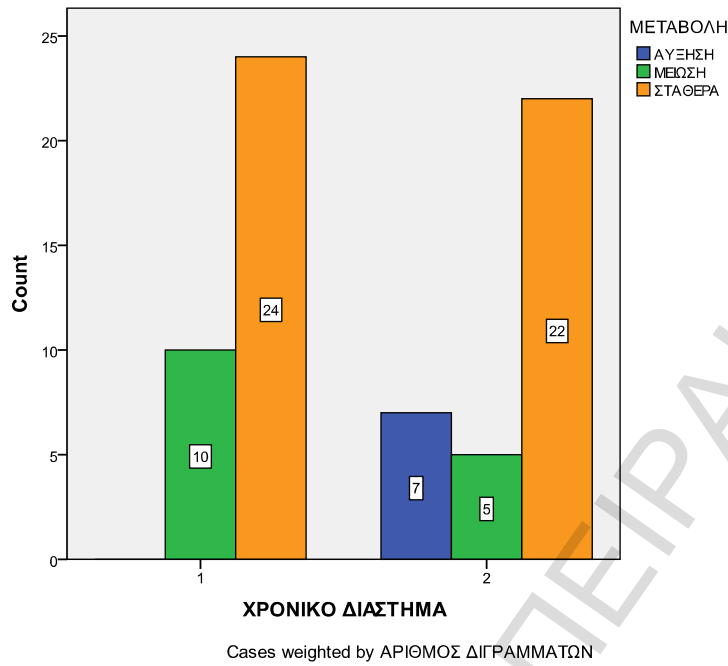
Τι συμβαίνει όμως όταν ο χρήστης πληκτρολογεί αντίθετα διγράμματα από τα οποία το ένα είναι συλλαβή όπως π.χ. ΑΚ και ΚΑ όπου οι χρόνοι κίνησης τους είναι ίσοι; Αν οι μέσοι χρόνοι πτήσης τους διαφέρουν αυτό προφανώς οφείλεται στον διαφορετικό χρόνο αναζήτησης. Για να διαπιστωθεί αν και κατά πόσο κάτι τέτοιο συμβαίνει πραγματοποιήθηκαν 20 t-test μεταξύ ζευγών τέτοιων αντίθετων διγραμμάτων. Τα αποτελέσματα έδειξαν ότι τα διγράμματα που είναι συλλαβές έχουν στατιστικά μικρότερους μέσους χρόνους πτήσης από τα αντίθετα τους διγράμματα, στους 19 από τους 20 στατιστικούς ελέγχους που πραγματοποιήθηκαν. Αν και τα δεδομένα δεν επαρκούσαν, ώστε να επεξεργαστούμε στατιστικά τη συμπεριφορά μεμονωμένα κάθε χρήστη που συμμετείχε, η σύγκριση των μέσων χρόνων πτήσης των διγραμμάτων αυτών έδειξε ότι οι μέσοι χρόνοι πτήσης κάθε συλλαβής ήταν τουλάχιστον ο μισός από αυτόν του αντίθετου διγράμματος και με πολύ μικρότερη τυπική απόκλιση. Το γεγονός αυτό φανερώνει ότι οι χρήστες έχουν αυτοματοποιήσει σε μεγαλύτερο βαθμό την πληκτρολόγηση συλλαβών.



Εικόνα 6.6 Γράφημα των αποτελεσμάτων των στατιστικών ελέγχων για την ισότητα του μ.ο. του χρόνου πτήσης των συλλαβών με τα αντίθετα διγράμματα τους

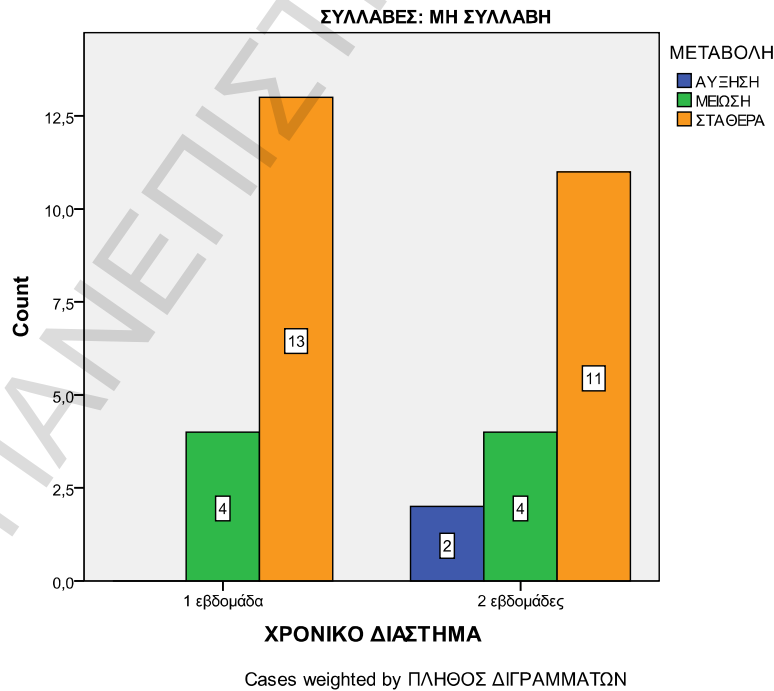
Για να ελέγξουμε αν διαφέρει ο χρόνος πτήσης κάθε συλλαβής ανάλογα με την λέξη στην οποία βρίσκεται πραγματοποιήσαμε αναλύσεις διακύμανσης κατά ένα παράγοντα (ANOVA). Δεν παρατηρήθηκε στατιστική διαφορά στους μέσους χρόνους πτήσης κάθε διγράμματος στις περισσότερες από τις λέξεις στις οποίες βρισκόταν. Κατά συνέπεια, ο μέσος χρόνος πτήσης κάθε διγράμματος μπορεί να υπολογιστεί από όλους τους χρόνους πτήσης του ανεξάρτητα από τις λέξεις στις οποίες περιέχεται. Οι χρόνοι πτήσης διαδοχικών γραμμάτων καθώς και οι χρόνοι πτήσης διαδοχικών διγραμμάτων βρέθηκαν να έχουν στατιστικά ασήμαντο συντελεστή γραμμικής συσχέτισης οπότε και θεωρούνται ανεξάρτητοι μεταξύ τους.

Για να εξεταστεί η χρονική σταθερότητα των χρόνων πτήσης των διγραμμάτων πραγματοποιήσαμε ανάλυση διακύμανσης ANOVA για 34 συλλαβές και 17 μη συλλαβές μεταξύ των 3 τεστ που πραγματοποιήθηκαν. Τα αποτελέσματα έδειξαν ότι μετά την παρέλευση μιας εβδομάδας οι χρόνοι πτήσης των 24 συλλαβών παρέμειναν σχεδόν αμετάβλητοι ενώ των υπόλοιπων 10 μειώθηκαν. Μετά από δύο εβδομάδες οι μέσοι χρόνοι πτήσης των 22 διατηρήθηκαν σταθεροί, των 5 μειώθηκαν και των υπόλοιπων 7 αυξήθηκαν. Αξίζει να σημειωθεί ότι οι συλλαβές των οποίων οι μέσοι χρόνοι πτήσης αυξήθηκαν έχουν χαμηλή συχνότητα εμφάνισης στην ελληνική γλώσσα.



Εικόνα 6.7 Γράφημα αναπαράστασης για το πώς μεταβλήθηκε ο μέσος χρόνος πτήσης των συλλαβών σε διάστημα 1 και 2 εβδομάδων

Όσον αφορά τα διγράμματα που δεν ήταν συλλαβές μετά την πρώτη εβδομάδα οι μέσοι χρόνοι πτήσης των 13 δεν σημείωσαν αλλαγή ενώ των 4 μειώθηκαν. Μετά από χρονικό διάστημα 2 εβδομάδων όταν και επαναλήφθηκε το 3^ο τεστ, οι μέσοι χρόνοι πτήσης παρέμειναν σταθεροί στα 11, μειώθηκαν στα 4 και αυξήθηκαν στα 2 διγράμματα. Η εξέταση των αντίστοιχων θηκογραμμάτων ήρθε να επιβεβαιώσει τις παραπάνω διαπιστώσεις.



Εικόνα 6.8 Γράφημα αναπαράστασης για το πώς μεταβλήθηκε ο μέσος χρόνος πτήσης των μη συλλαβών σε διαστήματα 1 και 2 εβδομάδων

Αντίθετα ο χρόνος πίεσης των γραμμάτων δείχνει να παρουσιάζει μικρότερη ενδο-ατομική μεταβλητότητα καθώς οι χρόνοι πίεσης κάθε γράμματος για κάθε έναν από τους 9 χρήστες που πραγματοποίησαν το τεστ παρουσίασαν μικρή διασπορά. Χαρακτηριστικό είναι ότι σε 108 αναλύσεις διακύμανσης ANOVA που πραγματοποιήσαμε στα 12 συχνότερα εμφανιζόμενα γράμματα των λέξεων, για τις 3 ημερομηνίες που πραγματοποιήθηκε το τεστ οι μέσοι χρόνοι πίεσης των 78 από τα 108 γράμματα (ποσοστό 72,2%) παρέμειναν αμετάβλητοι και στα 3 τεστ.

Τέλος, η ανάλυση διακύμανσης ANOVA μεταξύ των χρηστών έδειξε ότι η μέση τιμή του χρόνου πίεσης παρουσιάζει σημαντική δια-ατομική μεταβλητότητα καθώς μόνο σε ένα ζεύγος από τα 36 ζεύγη χρηστών που προκύπτουν από το συνδυασμό κάθε χρήστη με τους υπόλοιπους 8, οι μέσοι χρόνοι πίεσης δεν διέφεραν στατιστικά σημαντικά.

6.2 Αυθεντικοποίηση χρήστη με τον τρόπο πληκτρολόγησης

Για να εξετάσουμε την δυνατότητα επιβεβαίωσης ενός χρήστη με τον τρόπο πληκτρολόγησης εργαστήκαμε ως εξής:

Επειδή με την συνεχή εξάσκηση ο χρόνος πτήσης του κάθε διγράμματος για κάθε χρήστη βαίνει μειούμενος αλλά και για να αυξήσουμε την ομοιογένεια των δεδομένων αφαιρούμε το 10% των μεγαλύτερων χρόνων πτήσης κάθε διγράμματος, καθώς θεωρούμε ότι με την πάροδο του χρόνου ο χρήστης αυτοματοποιεί τον τρόπο πληκτρολόγησης με αποτέλεσμα ακραίες τιμές να εμφανίζονται όλο και πιο σπάνια. Η χρήση προσωπικού φίλτρου για την αφαίρεση ακραίων τιμών προτείνεται άλλωστε και στο [3] αλλά και σε προγενέστερες έρευνες από αυτή, ώστε να περιοριστεί η διακύμανση των χρόνων πτήσης κάθε διγράμματος.

Στη συνέχεια συγκρίνουμε τους χρόνους πτήσης των 3 πιο γρήγορων χρηστών (1,8 και 9) και υπολογίζουμε, από τους πίνακες συχνοτήτων, την πιθανότητα που υπάρχει, αν ένας από τους χρήστες έχει 100% πιθανότητα να πετυχαίνει χρόνους πτήσης εντός των ορίων που παρέχει το δείγμα του κάθε διγράμματος του (θεωρώντας ότι το δείγμα αυτό είναι αρκετά μεγαλύτερο ώστε η πιθανότητα ο χρήστης να σημειώσει χρόνους πτήσης εκτός αυτών των ορίων είναι πρακτικά πολύ μικρή), για τους 2 άλλους χρήστες να πληκτρολογήσουν το δίγραμμα μιμούμενοι τον τρίτο χρήστη.

Οι συγκρίσεις έγιναν σε 21 διγράμματα που ήταν συλλαβές και οι μέσες πιθανότητες λανθασμένης εισόδου (FAR) ήταν:

		USER		
		1	8	9
IMPOSTR	1	100,00	64,58	83,56
	8	48,61	100,00	70,91
	9	72,72	67,39	100,00

Και η μέση πιθανότητα αντιγραφής είναι 68%. Επειδή η πιθανότητα επιτυχημένης αντιγραφής της λέξης ισούται με το ποσοστό λανθασμένων επιβεβαιώσεων, δηλαδή:

$$FAR = P(word)$$

Αν θεωρήσουμε την πληκτρολόγηση διαφορετικών διγραμμάτων της λέξης ανεξάρτητα μεταξύ τους γεγονότα τότε:

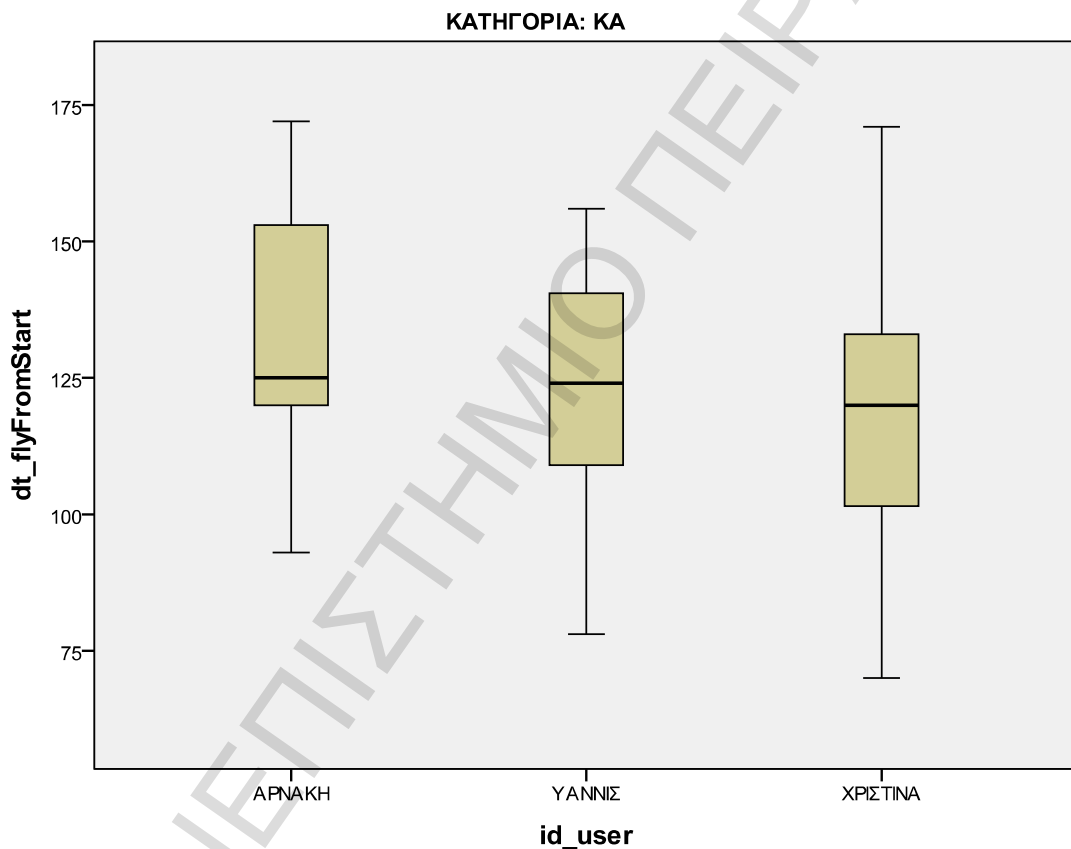
$$P(word) = (\bar{P}(digram))^v \text{ όπου } v: \text{ αριθμός συλλαβών λέξης}$$

αν θέλουμε το $FAR < 10\%$ θα πρέπει:

$$\log(P(word)) = \nu \log(\bar{P}(digram)) \Rightarrow \frac{\log(P(word))}{\log(\bar{P}(digram))} = \nu \approx 6$$

Αυτό σημαίνει ότι θα πρέπει η κωδική λέξη να έχει τουλάχιστον 6 συλλαβές ώστε η πιθανότητα να γίνει λανθασμένη επιβεβαίωση ενός εισβολέα να είναι μικρότερη του 10%. Δηλαδή η κωδική λέξη θα πρέπει να έχει τουλάχιστον 12 γράμματα. Η παρατήρηση αυτή συμφωνεί με τις μέχρι τώρα έρευνες [5] που προτείνουν ως ελάχιστο αριθμό γραμμάτων μιας κωδικής λέξης τα 9. Διγράμματα που είναι πολύ συχνά στη γλώσσα εμφανίζουν μεγαλύτερες πιθανότητες ταύτισης καθώς επίσης και αυτά που αποτελούνται από γράμματα που βρίσκονται στην ίδια στήλη του πληκτρολογίου.

Για παράδειγμα, τα θηκογράμματα και οι πιθανότητες για το δίγραμμα KA φαίνονται παρακάτω για τους 3 χρήστες:



Εικόνα 6.8 boxplot των χρόνων πτήσης του διγράμματος KA για κάθε ένα από τους 3 ταχύτερους χρήστες

		USER		
		APNAKH	YANNIS	XPISTINA
IMPOSTOR	APNAKH	100,00	80,00	90,00
	YANNIS	94,70	100,00	100,00
	XPISTINA	78,90	89,40	100,00

Τέλος μικρή τυπική απόκλιση ή μικρός μέσος χρόνος πτήσης δεν φαίνεται να μειώνουν τις πιθανότητες ταύτισης μεταξύ χρηστών ίδιου επιπέδου. Κατά συνέπεια είναι προτιμότερο να συγκρίνονται οι χρόνοι πτήσης διγραμμάτων που δεν είναι συλλαβές ή έχουν μικρή συχνότητα εμφάνισης στη γλώσσα, καθώς αυτές ο χρήστης τις έχει επαναλάβει αρκετές φορές κατά την πληκτρολόγηση των κωδικών σε αντίθεση με τους υπόλοιπους χρήστες που τις συναντούν σπάνια. Αυτό σημαίνει ότι ο συντελεστής βαρύτητας κάθε διγράμματος δεν πρέπει να εξαρτάται μόνο από τη διακύμανση των χρόνων πτήσης του [7], αλλά και από τη συχνότητα εμφάνισης του στην γλώσσα.

6.3 Αναγνώριση χρήστη με τον τρόπο πληκτρολόγησης

Για να εξεταστεί η δυνατότητα αναγνώρισης μέσω πληκτρολόγησης χρησιμοποιήθηκαν τα 10 διγράμματα που εμφανίζονταν συχνότερα στις λέξεις του τεστ. Το 1^ο τεστ χρησιμοποιήθηκε ως αρχείο εκπαίδευσης και το 2^ο τεστ ως κείμενο αναγνώρισης.

Για κάθε χρήστη, υπολογίστηκε η Ευκλείδεια απόσταση του μέσου χρόνου πτήσης κάθε διγράμματος με τον αντίστοιχο χρόνο του ίδιου διγράμματος αναφοράς και των 9 χρηστών.

$$D_{digramm_{ij}} = \sqrt{(t_{fly_2} - t_{fly_{ij}})^2}$$

όπου :

i: 1-9 (χρήστες) και

j:1-10 (διγράμματα)

Στη συνέχεια, οι αποστάσεις αθροίστηκαν και για τα δέκα διγράμματα κάθε χρήστη.

$$D_{ολικο_i} = \sum_{j=1}^{10} D_{digramm_{ij}}$$

Αν ο χρήστης εμφάνιζε την μικρότερη συνολική απόσταση με το δικό του κείμενο αναφοράς τότε η αναγνώριση ήταν επιτυχής. Με τον τρόπο αυτό οι 4 στους 9 χρήστες αναγνωρίστηκαν επιτυχώς.

Ανάλογη διαδικασία εφαρμόστηκε μεταξύ του 2^{ου} και 3^{ου} τεστ όπου το 2^ο ήταν αυτή τη φορά το κείμενο αναφοράς και το 3^ο το κείμενο αναγνώρισης. Τα αποτελέσματα και σε αυτή την περίπτωση ήταν 4 επιτυχείς αναγνωρίσεις.

Το ποσοστό αναγνώρισης 44% φαντάζει μικρό αλλά αν αναλογιστούμε ότι η πιθανότητα αναγνώρισης με βάση την τυχαία επιλογή χρήστη είναι περίπου 11%, ένα τέτοιο σύστημα αναγνώρισης τετραπλασιάζει τις πιθανότητες επιτυχούς αναγνώρισης ενός χρήστη.

Για να εξετάσουμε αν ο χρόνος πίεσης παρουσιάζει ικανή δια-ατομική μεταβλητότητα ώστε να διακρίνουμε τα άτομα μεταξύ τους με βάση το χαρακτηριστικό αυτό, επαναλάβουμε την ίδια διαδικασία που πραγματοποιήσαμε και για τα διγράμματα. Δηλαδή επιλέξαμε τα 10 πιο συχνά εμφανιζόμενα γράμματα στις λέξεις του τεστ και για κάθε χρήστη υπολογίστηκε η Ευκλείδεια απόσταση μεταξύ του μέσου χρόνου πίεσης κάθε γράμματος και καθενός εκ των 9 μέσων χρόνων πίεσης του, που σημείωσαν οι χρήστες στο αμέσως προηγούμενο τεστ.

Στη συνέχεια, οι αποστάσεις αθροίστηκαν και για τα δέκα γράμματα κάθε χρήστη. Αν ο χρήστης εμφάνιζε την μικρότερη συνολική απόσταση με το δικό του κείμενο αναφοράς τότε η αναγνώριση ήταν επιτυχής. Με τον τρόπο αυτό ήταν επιτυχείς οι 14 από τις 18 προσπάθειες αναγνώρισης.

6.4 Συμπεράσματα - Περίληψη

Από τους στατιστικούς ελέγχους φαίνεται ότι οι συλλαβές πληκτρολογούνται ταχύτερα από τα αντίθετα τους διγράμματα σε ποσοστό 95%, ακόμα και σε συλλαβές που έχουν το παρόμοιο ποσοστό εμφάνισης στη γλώσσα με τα αντίθετα διγράμματα τους. Επειδή ο χρόνος κίνησης κατά την πληκτρολόγηση κάθε συλλαβής είναι ίδιος με το χρόνο κίνησης κατά την πληκτρολόγηση του αντίθετου διγράμματος, συμπεραίνουμε ότι η διαφορά που προκύπτει οφείλεται στον διαφορετικό χρόνο αναζήτησης κατά την πληκτρολόγηση τους. Το συμπέρασμα αυτό ενισχύουν και τα αποτελέσματα που προκύπτουν από τις γραμμικές παλινδρομήσεις όπου φαίνεται ότι ο μέσος χρόνος πτήσης κάθε διγράμματος μειώνεται όσο αυξάνεται η συχνότητα εμφάνισης στη γλώσσα του 2^{ου} γράμματος του και όταν το δίγραμμο είναι συλλαβή. Αυτό σημαίνει ότι συλλαβές που εμφανίζονται πολύ συχνά στη γλώσσα έχουν μικρότερο χρόνο αναζήτησης και κατά συνέπεια εμφανίζουν μεγαλύτερη χρονική σταθερότητα, πληκτρολογούνται αντανακλαστικά και επηρεάζονται λιγότερο από τη συναισθηματική, ψυχολογική και πνευματική κατάσταση του χρήστη.

Άρα διανύσματα τέτοιων συλλαβών μπορούν να χρησιμοποιηθούν ως δείκτες σύγκρισης με διανύσματα ασταθών διγραμμάτων και οι αποκλίσεις που σημειώνονται να αξιολογούν την τρέχουσα κατάσταση του χρήστη. Περαιτέρω έρευνα απαιτείται για να διαπιστωθεί το αν και κατά πόσο μεταβάλλονται αυτοί οι χρόνοι πτήσης όταν τροποποιηθεί η πνευματική, συναισθηματική ή ψυχική κατάσταση του χρήστη όπως πχ συμβαίνει με την ύπαρξη θορύβου ή άλλων ερεθισμάτων.

Όπως είδαμε υπάρχει σχέση μεταξύ του μέσου χρόνου πτήσης των συλλαβών και των μη συλλαβών, η οποία παραμένει σταθερή για μικρά τουλάχιστον χρονικά διαστήματα. Η σχέση αυτή θα μπορούσε να χρησιμοποιηθεί για την ταξινόμηση των χρηστών σε διαφορετικά επίπεδα εμπειρίας και στη χρήση στερεοτύπων.

Όπως είδαμε οι συχνότερα εμφανιζόμενες συλλαβές ενώ παρουσιάζουν μικρούς χρόνους πτήσης και μικρές τιμές τυπικής απόκλισης, ειδικά σε έμπειρους χρήστες, παρουσιάζουν μικρή δια-ατομική μεταβλητότητα και δεν πρέπει να χρησιμοποιούνται ως κριτήριο αυθεντικοποίησης ή αναγνώρισης χρήστη. Τα διανύσματα διγραμμάτων που χρησιμοποιούνται για το σκοπό αυτό θα πρέπει να αποτελούνται από διγράμματα με μικρή συχνότητα εμφάνισης και μικρή διασπορά. Κατά συνέπεια είναι προτιμότερο να συγκρίνονται οι χρόνοι πτήσης διγραμμάτων που δεν είναι συλλαβές ή έχουν μικρή συχνότητα εμφάνισης στη γλώσσα, καθώς αυτές ο χρήστης τις έχει επαναλάβει αρκετές φορές κατά την πληκτρολόγηση των κωδίκων σε αντίθεση με τους υπόλοιπους χρήστες που τις συναντούν σπάνια.

Τέλος οι χρόνοι πτήσης των γραμμάτων φαίνεται να παρουσιάζουν μικρότερη ενδο-ατομική μεταβλητότητα ακόμα και σε άπειρους χρήστες. Το ποσοστό επιτυχούς αναγνώρισης χρήστη μόνο με τη χρήση των χρόνων πτήσης άγγιξε το 77% σε αντίθεση με το πενιχρό 44% το οποίο επιτεύχθηκε με τη χρήση των χρόνων πτήσης των διγραμμάτων.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] R. Joyce and G. Gupta, "Identity authentication based on keystroke latencies," Commun. ACM, vol. 33, no. 2, pp. 168–176, 1990.
- [2] Monroe, F., Rubin, A., "Authentication via Keystroke Dynamics", Proceedings of the 4th ACM Conference on Computer and Communications Security, p 48-56, April 1997
- [3] F. Monroe and A. D. Rubin, "Keystroke dynamics as a biometric for authentication," Future Gen. Comput. Syst., vol. 16, no. 4, pp. 351–359, 2000.
- [4] Daniele Gunetti, Claudia Picardi, and Giancarlo Ruffo "Dealing with Different Languages and Old Profiles in Keystroke Analysis of Free Text", S. Bandini and S. Manzoni (Eds.): AI*IA 2005, LNAI 3673, pp. 347–358, 2005.
- [5] Cheng-Huang Jiang, Shiuhyng Shieh, Jen-Chien Liu "Keystroke Statistical Learning Model for Web Authentication", ASIACCS'07, March 20–22, 2007, Singapore
- [6] Gaines, R., Lisowski, W., Press, S., and Shapiro, N. 1980. Authentication by keystroke timing: Some preliminary results. Rand. Report R-256-NSF. Rand Corporation.
- [7] Σιατρας Σ., "Σύστημα αυθεντικοποίησης μέσω αναγνώρισης τρόπου πληκτρολόγησης", Διπλωματική εργασία, ΑΠΘ, 2004
- [8] Ι. Μανωλόπουλος, Α.Ν. Παπαδόπουλος "Συστήματα Βάσεων Δεδομένων: θεωρία και πρακτική εφαρμογή", ISBN: 1-85233-977-2, 2006.
- [9] Δ. Καραγεωργός "Στατιστική, Περιγραφική & Επαγωγική", ISBN: 960-460-623-9, 2001
- [10] Susan B. Gerber, Kristin Voelkl Finn "Using SPSS for Windows: Data Analysis and Graphics", ISBN: 0-387-40083-4, 2005.
- [11] B. Eckel "Thinking in Java" ISBN: 0-131-87248-6, 2006
- [12] W. Boggs, M. Boggs "Mastering UML with Rational Rose 2002", ISBN: 0–7821–4017–3, 2002.