



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Ανάλυση και Αξιολόγηση Επιχειρησιακών Διεργασιών στην Τοπική
Αυτοδιοίκηση με χρήση Τεχνικών Εξόρυξης Διαδικασιών**
Βασίλειος Χριστοδούλου
711161028



Επιβλέπων καθηγητής: Αλέξανδρος Μπουσδέκης
Συνεπιβλέπων καθηγητής: Γεώργιος Μιαούλης
Διπλωματική εργασία υποβληθείσα στο Τμήμα Μηχανικών Πληροφορικής και Υπολογιστών του
Πανεπιστημίου Δυτικής Αττικής
Αθήνα, Χειμερινό εξάμηνο 2021-2022





ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Ανάλυση και Αξιολόγηση Επιχειρησιακών Διεργασιών στην Τοπική Αυτοδιοίκηση με χρήση Τεχνικών Εξόρυξης Διαδικασιών

Μέλη Εξεταστικής Επιτροπής

1. Γεώργιος Μιαούλης
2. Δημήτριος Μάγος
3. Γεώργιος Μπαρδής

Η πτυχιακή/διπλωματική εργασία εξετάστηκε επιτυχώς από την κάτωθι Εξεταστική Επιτροπή:

α/α	Όνοματεπώνυμο	Ιδιότητα	Ψηφιακή Υπογραφή
1	Γεώργιος Μιαούλης	ΔΕΠ	
2	Δημήτριος Μάγος	ΔΕΠ	
3	Γεώργιος Μπαρδής	ΔΕΠ	

ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

Ο κάτωθι υπογεγραμμένος ΒΑΣΙΛΕΙΟΣ ΧΡΙΣΤΟΔΟΥΛΟΥ του ΧΡΗΣΤΟΥ, με αριθμό μητρώου 711161028 φοιτητής του Τμήματος Μηχανικών Πληροφορικής και Υπολογιστών της Σχολής Μηχανικών του Πανεπιστημίου Δυτικής Αττικής, δηλώνω ότι: «Είμαι συγγραφέας αυτής της πτυχιακής/διπλωματικής εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της, είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, οι όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε ακριβώς είτε παραφρασμένες, αναφέρονται στο σύνολό τους, με πλήρη αναφορά στους συγγραφείς, τον εκδοτικό οίκο ή το περιοδικό, συμπεριλαμβανομένων και των πηγών που ενδεχομένως χρησιμοποιήθηκαν από το διαδίκτυο. Επίσης, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από μένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου, όσο και του Ιδρύματος.

Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του πτυχίου μου».

Ο/Η Δηλών/ούσα
Βασίλειος Χριστοδούλου



ΕΥΧΑΡΙΣΤΙΕΣ

Για την εκπόνηση της παρούσας πτυχιακής/διπλωματικής εργασίας, θα ήθελα να ευχαριστήσω τον καθηγητή μου κ. Αλέξανδρο Μπουσδέκη, για τις συμβουλές του και την υποστήριξή του σε όλη την διάρκεια. Ακόμα, θα ήθελα να ευχαριστήσω την οικογένεια μου για την στήριξή τους και την υπομονή τους καθ' όλη την διάρκεια των σπουδών μου.

ΠΕΡΙΛΗΨΗ

Στην σημερινή εποχή μία επιχείρηση/οργανισμός χρειάζεται να διαχειριστεί μεγάλους όγκους δεδομένων ,να εκτελέσει καθημερινά απεριόριστες εργασίες και να διεκπεραιώσει ζητήματα που προκύπτουν ασταμάτητα.Η ικανότητα μίας επιχείρησης να ανταπεξέλθει σε αυτά τα θέματα είναι μείζονος σημασίας για την βιωσιμότητα της.Αν σε κάποιες περιπτώσεις η απόδοση και η παραγωγικότητα της υστερεί,τότε δημιουργείται πρόβλημα στην επιχείρηση ή στον οργανισμό.Η επιχείρηση αποσκοπεί ,μέσα από τις εργασίες να αυξήσει τα έσοδα,ενώ ο οργανισμός στοχεύει στην επίλυση αιτημάτων που προκύπτουν.Η εξόρυξη διαδικασιών και τα εργαλεία της μπορούν να μας βοηθήσουν να παρακολουθήσουμε τις εργασίες και να βγάλουμε συμπεράσματα για αυτές και τον τρόπο που εκτελούνται.Οι τεχνικές και οι πλατφόρμες της εξόρυξης μας δίνουν την δυνατότητα να αναλύσουμε τις εργασίες και την συμπεριφορά τους,να εντοπίσουμε προβλήματα και ευπάθειες σε ένα σύστημα και να περιορίσουμε ή εξαλείψουμε τα προβλήματα.

Η εφαρμογή της εξόρυξης διαδικασιών σε δραστηριότητες και εργασίες είναι ένα εργαλείο για να μελετήσουμε την συμπεριφορά και τα μοτίβα που αναπτύσσονται,να ανακαλύψουμε ανωμαλίες και να βελτιώσουμε τον τρόπο εκτέλεσης των εργασιών.Με αυτόν τον τρόπο μία επιχείρηση θα μπορεί να αυξήσει την απόδοση της και να καλυτερέψει τον τρόπο λειτουργία της,για να στηρίξει την βιωσιμότητα της.Με άμεσο αντίκτυπο και στους ίδιους τους εργαζομένους,όπου θα μπορούν να βελτιώσουν το βιοτικό τους επίπεδο.

Στην διπλωματική εργασία χρησιμοποιούμε ένα σύνολο δεδομένων από το 11th International Workshop on Business Process Intelligence 2015,Business Processing Intelligence Challenge (BPIC),που περιλαμβάνει δεδομένα από 5 Ολλανδικούς δήμους.Τα στοιχεία περιέχουν εργασίες που έλαβαν χώρα από την τοπική αυτοδιοίκηση ,σε διάστημα μίας 4ετίας.Περιλαμβάνονται πολλές διαφορετικές εργασίες,που υποδηλώνονται και με δύο κωδικούς (χαρακτηριστικό:όνομα) και με ετικέτες, τόσο στα ολλανδικά (χαρακτηριστικό taskNameNL) όσο και στα αγγλικά (χαρακτηριστικό taskNameEN).Οι υποθέσεις περιέχουν εργασίες σε πολλά διαφορετικά στάδια της ανάπτυξης τους,καθώς και τον υπεύθυνο που τις είχε αναλάβει και το εκτιμώμενο κόστος.Οι διαδικασίες στους 5 δήμους πρέπει να είναι πανομοιότυπες,αλλά μπορούν να διαφέρουν ελάχιστα μεταξύ τους.

Με την χρήση του open source λογισμικού Prom Tools για εξόρυξη διαδικασιών,που χρησιμοποιείται ευρέως στον τομέα της εξόρυξης,θα είμαστε σε θέση για κάθε δήμο να αναλύσουμε τις διαδικασίες και να τις μελετήσουμε.Στη συνέχεια θα μπορούμε να εντοπίσουμε σημεία ενδιαφέροντος,όπως ανομοιογένειες,παθογένειες και σφάλματα σε κάθε δήμο,και να τα επεξεργαστούμε περαιτέρω με φιλτράρισμα και με αλγόριθμους mining.Με αυτήν την τεχνική θα μπορούμε να ανακαλύψουμε ευπάθειες σε κάθε δήμο που μειώνουν την απόδοση και να τις εξαλείψουμε.Καταλήγοντας σε μία κατάσταση όπου θα μπορούμε να εκτελέσουμε εργασίες με τον πιο αποδοτικό τρόπο,αυξάνοντας κατά πολύ την παραγωγικότητα που αποζητείται.

Keywords:Εξόρυξη διεργασιών,Process mining,Τοπική αυτοδιοίκηση,Αρχείο καταγραφής συμβάντων,Ανάλυση και αξιολόγηση,Επιχειρησιακές διεργασίες

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΚΕΦΑΛΑΙΟ 1

- 1.1 ΕΙΣΑΓΩΓΗ.....9
- 1.2 ΑΝΤΙΚΕΙΜΕΝΟ ΤΗΣ ΔΙΠΛΩΜΑΤΙΚΗΣ.....11
- 1.3 ΔΙΑΡΘΡΩΣΗ ΤΗΣ ΔΙΠΛΩΜΑΤΙΚΗΣ.....11

ΚΕΦΑΛΑΙΟ 2

- 2.1 ΕΠΙΧΕΙΡΗΜΑΤΙΚΗ ΕΥΦΕΙΑ (BUSINESS INTELLIGENCE).....13
- 2.2 ΔΙΑΧΕΙΡΙΣΗ ΕΠΙΧΕΙΡΗΜΑΤΙΚΩΝ ΔΙΑΔΙΚΑΣΙΩΝ (BPM).....15
- 2.3 ΕΞΟΡΥΞΗ ΔΕΔΟΜΕΝΩΝ.....16
- 2.4 ΕΞΟΡΥΞΗ ΔΙΑΔΙΚΑΣΙΩΝ.....17
 - 2.4.1 ΕΙΣΑΓΩΓΗ.....17
 - 2.4.2 ΙΔΙΟΤΗΤΕΣ ΚΑΙ ΕΦΑΡΜΟΓΕΣ.....18
 - 2.4.3 ΕΡΓΑΛΕΙΑ ΕΞΟΡΥΞΗΣ ΔΙΑΔΙΚΑΣΙΩΝ.....21
- 2.5 PROCESS MODELS ΚΑΙ MINING ALGORITHMS.....23
- 2.6 ΑΡΧΕΙΑ ΚΑΤΑΓΡΑΦΗΣ ΣΥΜΒΑΝΤΩΝ.....26
 - 2.6.1 ΔΟΜΗ ΑΡΧΕΙΩΝ.....26
 - 2.6.2 ΠΡΟΤΥΠΑ ΑΡΧΕΙΩΝ.....28
- 2.7 ΠΡΟΚΛΗΣΕΙΣ ΣΤΗΝ ΤΟΠΙΚΗ ΑΥΤΟΔΙΟΙΚΗΣΗ.....32

ΚΕΦΑΛΑΙΟ 3

- 3.1 ΜΕΘΟΔΟΛΟΓΙΑ ΕΞΟΡΥΞΗΣ ΔΙΑΔΙΚΑΣΙΩΝ.....35
- 3.2 ΕΞΟΡΥΞΗ ΔΙΑΔΙΚΑΣΙΩΝ ΜΕ ΤΟ PROM TOOLS.....37
 - 3.2.1 ΕΙΣΑΓΩΓΗ ΑΡΧΕΙΩΝ ΚΑΤΑΓΡΑΦΗΣ ΣΥΜΒΑΝΤΩΝ.....37
 - 3.2.2 EVENT LOG VISUALIZATION.....39
 - 3.2.3 PROCESS MODEL.....50
 - 3.2.4 FILTERING EVENT LOG.....53

ΚΕΦΑΛΑΙΟ 4

- 4.1 ΜΕΛΕΤΗ ΤΗΣ ΔΙΠΛΩΜΑΤΙΚΗΣ.....57
 - 4.1.1 ΕΙΣΑΓΩΓΗ ΑΡΧΕΙΩΝ ΚΑΤΑΓΡΑΦΗΣ ΣΥΜΒΑΝΤΩΝ.....58
 - 4.1.2 EVENT LOG VISUALIZATION.....66
 - 4.1.3 PROCESS MODEL.....79
 - 4.1.4 FILTERING EVENT LOG.....83

ΚΕΦΑΛΑΙΟ 5

- ΣΥΜΠΕΡΑΣΜΑ.....111

ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ

ΛΙΣΤΑ ΕΙΚΟΝΩΝ

EIKONA 1 Business intelligence in architecture and technology.....	13
EIKONA 2 Η πορεία του Process mining.....	14
EIKONA 3 Σύνδεση ανάμεσα σε εξόρυξη δεδομένων και διαχείριση επιχειρηματικών διαδικασιών.....	16
EIKONA 4 Λειτουργία του process mining στην καθημερινότητα.....	20
EIKONA 5 Κύριοι δημιουργοί λογισμικών εξόρυξης διαδικασιών ανά τον κόσμο.....	21
EIKONA 6 Μερίδιο αγοράς προμηθευτών λογισμικού εξόρυξης διεργασιών.....	22
EIKONA 7 Η αγορά της ανάλυσης διεργασιών.....	23
EIKONA 8 Δομή διαγράμματος ροής ή μοντέλου διαδικασιών.....	24
EIKONA 9 Σύγκριση μεταξύ των mining αλγόριθμων.....	25
EIKONA 10 Παράδειγμα από αρχείο καταγραφής συμβάντων με ίχνη και διαδικασίες, όπου το καθένα περιλαμβάνει attributes.....	26
EIKONA 11 Εικόνα από Event Log Explorer με συμβάντα στο λειτουργικό σύστημα Microsoft Windows.....	28
EIKONA 12 Δομή ενός XES αρχείου.....	29
EIKONA 13 Δομή ενός ίχνους με διαδικασίες και attributes.....	29
EIKONA 14 Δομή των log file του παραδείγματος μας.....	30
EIKONA 15 Λίστα με όλα τα attributes και την ερμηνεία τους του παραδείγματος μας.....	31
EIKONA 16 Το περιβάλλον του Prom Tools.....	38
EIKONA 17 Επιλογή για εισαγωγή αρχείου στο λογισμικό.....	38
EIKONA 18 Επιλογή του επιθυμητού αρχείου από το σύστημα μας.....	39
EIKONA 19 Αρχική σελίδα.....	39
EIKONA 20 Λίστα με τις διαθέσιμες τεχνικές εικονικοποίησης των δεδομένων.....	40
EIKONA 21 Λίστα με τα αρχεία και των διαθέσιμων ενεργειών.....	40
EIKONA 22 Απεικόνιση των στοιχείων του αρχείου καταγραφής συμβάντων.....	41
EIKONA 23 Καρτέλα browser στον inspector.....	42
EIKONA 24 Καρτέλα explorer στον inspector.....	43
EIKONA 25 Καρτέλα explorer-στοιχεία για το ίχνη.....	43
EIKONA 26 Καρτέλα Log Attributes.....	44
EIKONA 27 Καρτέλα Log Summary.....	45
EIKONA 28 Explore Event Log διάγραμμα.....	47
EIKONA 29 Αναλυτικές λεπτομέρειες των υποθέσεων.....	47
EIKONA 30 Πληροφορίες για το αρχείο.....	48
EIKONA 31 Περιβάλλον Dotted Chart.....	48
EIKONA 32 Παραμετροποίηση Dotted Chart.....	49
EIKONA 33 Παράδειγμα Dotted Chart.....	49
EIKONA 34 Παράδειγμα Alpha miner.....	51
EIKONA 35 Παράδειγμα Heuristic miner.....	51
EIKONA 36 Παράδειγμα Inductive Visual miner.....	52
EIKONA 37 Παράδειγμα Fuzzy miner.....	53
EIKONA 38 Καρτέλα Actions.....	55
EIKONA 39 Dashboard 1ου δήμου.....	59
EIKONA 40 Dashboard 2ου δήμου.....	60
EIKONA 41 Dashboard 3ου δήμου.....	61
EIKONA 42 Dashboard 4ου δήμου.....	62
EIKONA 43 Dashboard 5ου δήμου.....	63
EIKONA 44 Ανάλυση κωδικών διαδικασιών.....	66

EIKONA 45 Καρτέλα browser του inspector για τους δήμους της μελέτης.....	67
EIKONA 46 Καρτέλα explorer του inspector για τους δήμους της μελέτης.....	67
EIKONA 47 Καρτέλα Log Attributes για τη δομή των αρχείων της μελέτης.....	68
EIKONA 48 Καρτέλα Log Summary των 5 δήμων.....	70
EIKONA 49 Explore Event Log 1ου δήμου.....	72
EIKONA 50 Explore Event Log 2ου δήμου.....	73
EIKONA 51 Explore Event Log 3ου δήμου.....	74
EIKONA 52 Explore Event Log 4ου δήμου.....	75
EIKONA 53 Explore Event Log 5ου δήμου.....	76
EIKONA 54 Dotted Chart των δήμων.....	79
EIKONA 55 Alpha miner plug in.....	80
EIKONA 56 Heuristic miner plug in.....	81
EIKONA 57 Fuzzy miner plug in.....	82
EIKONA 58 Inductive Visual miner plug in.....	82
EIKONA 59 Καρτέλα Actions για επιλογή plug in.....	83
EIKONA 60 Filter Log By attributes plug in.....	84
EIKONA 61 Dashboard και Process model 1ου αρχείου για εργασία 01_HOOFD_100.....	85
EIKONA 62 Dashboard και Process model 2ου αρχείου για εργασία 01_HOOFD_100.....	87
EIKONA 63 Dashboard και Process model 3ου αρχείου για εργασία 01_HOOFD_100.....	88
EIKONA 64 Dashboard και Process model 4ου αρχείου για εργασία 01_HOOFD_100.....	90
EIKONA 65 Dashboard και Process model 5ου αρχείου για εργασία 01_HOOFD_180.....	92
EIKONA 66 Dashboard και Process model 1ου αρχείου για ίχνη με λιγότερες από 5 ενέργειες.....	93
EIKONA 67 Dashboard και Process model 2ου αρχείου για ίχνη με λιγότερες από 5 ενέργειες.....	94
EIKONA 68 Dashboard και Process model 3ου αρχείου για ίχνη με λιγότερες από 5 ενέργειες.....	95
EIKONA 69 Dashboard και Process model 4ου αρχείου για ίχνη με λιγότερες από 5 ενέργειες.....	96
EIKONA 70 Dashboard και Process model 5ου αρχείου για ίχνη με λιγότερες από 5 ενέργειες.....	97
EIKONA 71 Filter Log by trace length.....	98
EIKONA 72 Dashboard και Dotted Chart 1ου αρχείου για ίχνη με περισσότερες από 80 ενέργειες.....	99
EIKONA 73 Dashboard και Dotted Chart 2ου αρχείου για ίχνη με περισσότερες από 80 ενέργειες.....	100
EIKONA 74 Dashboard και Dotted Chart 3ου αρχείου για ίχνη με περισσότερες από 80 ενέργειες.....	101
EIKONA 75 Dashboard και Dotted Chart 4ου αρχείου για ίχνη με περισσότερες από 80 ενέργειες.....	102
EIKONA 76 Dashboard και Dotted Chart 5ου αρχείου για ίχνη με περισσότερες από 80 ενέργειες.....	103
EIKONA 77 Generate log from org:resource perspective για τον 1ο δήμο.....	105
EIKONA 78 Generate log from org:resource perspective για τον 2ο δήμο.....	105
EIKONA 79 Generate log from org:resource perspective για τον 3ο δήμο.....	106
EIKONA 80 Generate log from org:resource perspective για τον 4ο δήμο.....	106
EIKONA 81 Generate log from org:resource perspective για τον 5ο δήμο.....	107

ΚΕΦΑΛΑΙΟ 1

1.1 ΕΙΣΑΓΩΓΗ

ΕΠΙΧΕΙΡΗΣΕΙΣ ΚΑΙ ΟΡΓΑΝΙΣΜΟΙ

Οι επιχειρήσεις και οι οργανισμοί έχουν ως κύριο σκοπό της λειτουργίας τους, το κέρδος προσφέροντας στους καταναλωτές αγαθά ή υπηρεσίες. Με τον όρο επιχείρηση εξάλλου εννοούμε την ποριστική οικονομική μονάδα που αποτελεί αυτοτελή και υπεύθυνη οργάνωση παραγωγικών συντελεστών και διαχείρισης συναλλαγών με τις οποίες και επιδιώκει το μέγιστο δυνατό κέρδος.

Το δε κέρδος κατά κανόνα θα πρέπει να υπερβαίνει την αντίστοιχη συνήθη αμοιβή (ως αντιμισθία) της διοικητικής ή εκτελεστικής εργασίας που επιτελείται σ' αυτήν. Έτσι με τον παραπάνω ορισμό δίδεται σαφώς ως κύριο κριτήριο το κέρδος, του οποίου και ξεχωρίζει από κάποια άλλη οικονομική μονάδα, αφού όλα τα προαναφερθέντα χαρακτηριστικά είναι κοινά και αφηρημένα, με συνέπεια να μη λαμβάνονται ως κριτήρια διάκρισης μεταξύ των δύο εννοιών. Οι επιχειρήσεις και οι οργανισμοί είναι γενικά παραγωγικές μονάδες με διάφορες νομικές μορφές. Σε αυτές, ένα ή περισσότερα άτομα αποφασίζουν για την παραγωγή. Η επιδίωξη του μέγιστου κέρδους αποτελεί βασικό χαρακτηριστικό τους, διότι όσο μεγαλύτερο είναι το κέρδος, τόσο πιο βέβαια είναι και η επιβίωση και η ανάπτυξη τους μακροχρόνια. [1]

Όσων αφορά τους οργανισμούς, εκείνοι δεν εντάσσονται τόσο στο κερδοσκοπικό κομμάτι που λειτουργούν οι επιχειρήσεις, αλλά στοχεύουν στη χρήση νόμιμων μέσων και ενεργειών που αποσκοπούν στην επίτευξη συγκεκριμένου αποτελέσματος για την ικανοποίηση του γενικού συμφέροντος των πολιτών ενός κράτους.

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΚΑΙ ΕΙΔΗ ΕΠΙΧΕΙΡΗΣΕΩΝ

Κύρια χαρακτηριστικά γνωρίσματα μιας επιχείρησης, που αποτελούν και τις αναγκαίες προϋποθέσεις, είναι τα ακόλουθα:

1. Οικονομική μονάδα που να προϋποθέτει μόνιμο συνδυασμό συντελεστών παραγωγής.
2. Οικονομική μονάδα αυτοτελής.
3. Η παραγωγή της να απευθύνεται σε άγνωστο καταναλωτικό κοινό και τέλος
4. Να μην είναι πρωτογενής, αλλά να διαμορφώνεται σε ορισμένο κοινωνικοοικονομικό περιβάλλον χαρακτηριζόμενο τόσο εις χρήμα αποτίμηση, όσο και στην επιδίωξη κέρδους

Γενικά οι επιχειρήσεις διακρίνονται ανάλογα εκ του αντικειμένου της δράσης τους, του ιδιοκτησιακού επιχειρηματικού κεφαλαίου, εκ της νομικής μορφής τους και κατά μέγεθος:

Επιχειρήσεις εξ αντικειμένου δράσης

1. Επιχειρήσεις πρωτογενούς παραγωγής.
2. Επιχειρήσεις παραγωγής ή μεταποίησης ή μετασχηματισμού. Στη κατηγορία αυτή εντάσσονται όλες οι βιομηχανίες και βιοτεχνίες.
3. Επιχειρήσεις γενικού εμπορίου, ή εμπορικές.
4. Επιχειρήσεις ασφαλιστικές.
5. Επιχειρήσεις παροχή υπηρεσιών
6. Τραπεζικές επιχειρήσεις.
7. Επιχειρήσεις μεταφορών, που διακρίνονται σε χερσαίες, θαλάσσιες και εναέριας.

Επιχειρήσεις εκ του φορέα τους

1. Ιδιωτικές επιχειρήσεις
2. Δημόσιες επιχειρήσεις
3. Μικτές επιχειρήσεις

Στους οργανισμούς που αναφερόμαστε εντάσσονται κυρίως κρατικές, αυτοδιοικητικές τις περισσότερες φορές, μονάδες όπου στοχεύουν στην εξυπηρέτηση αναγκών των φυσικών και νομικών προσώπων και γενικά την άσκηση της δημόσιας διοίκησης. Επισημαίνουμε εδώ ότι η δημόσια διοίκηση ενός κράτους αποτελεί μέρος του κρατικού μηχανισμού και είναι απόλυτα συνυφασμένη με την κυβέρνηση. Στο συγκεκριμένο παράδειγμα της εργασίας ασχολούμαστε με τους δήμους, μία οντότητα της δημόσιας διοίκησης, όπου αποτελούν μέρος της τοπικής αυτοδιοίκησης ενός κράτους. Είναι μονάδες ανεξάρτητες έως ένα βαθμό, που περιλαμβάνουν μία πόλη ή κωμόπολη και τα γύρω χωριά της και μερικές φορές περιλαμβάνει μόνο κάποια συνοικία ή προάστιο μιας μεγαλούπολης, ή και αντίθετα μόνο χωριά μιας απομονωμένης ορεινής περιοχής ή ενός νησιού. Διευκρινίζουμε εδώ ότι στην Ελλάδα ο δήμος αποτελεί *πρωτόβαθμιο* Οργανισμό Τοπικής Αυτοδιοίκησης (ΟΤΑ). Σε άλλες χώρες οι α' βαθμιοί ΟΤΑ αναφέρονται αδιακρίτως ως κοινότητες. [2]

Στις παραπάνω περιπτώσεις που αναφέραμε είτε πρόκειται για επιχείρηση είτε για οργανισμό αυτές οι μονάδες αποσκοπούν στην υψηλή παραγωγικότητα και απόδοση, με σκοπό είτε το κέρδος είτε την εξυπηρέτηση αναγκών των ατόμων αντίστοιχα. Στις περιπτώσεις των οργανισμών και ειδικά των δήμων στόχος είναι γενικά η εύρυθμη λειτουργία της κοινότητας των κατοίκων και η προώθηση των συμφερόντων αυτών. Σε κάθε κράτος οι εκάστοτε κυβερνήσεις χαράσσουν τις κατευθυντήριες γραμμές και οι διοικήσεις τις εφαρμόζουν.

Με την ανάπτυξη της επιστήμης και της τεχνολογίας που έχουμε στην διάθεση μας την σημερινή εποχή, είμαστε σε θέση να αντιμετωπίσουμε αυτά τα ζητήματα και κάθε επιχείρηση να μπορεί να αναπτύξει την παραγωγή της και την διάθεση στοχευμένων προϊόντων και υπηρεσιών, μέσω και του διαδικτύου. Αξίζει να σημειωθεί ότι τα τελευταία χρόνια, πολλές επιχειρήσεις έχουν επενδύσει στην ραγδαία ανάπτυξη του διαδικτύου για να προσφέρουν τις υπηρεσίες τους. Ακόμα με την χρήση της

εξόρυξης διαδικασιών (process mining) μπορούμε να αναλύσουμε τις καθημερινές εργασίες που λαμβάνουν χώρα σε κάθε επιχείρηση ή οργανισμό. Με αυτόν τον τρόπο αναλύοντας σε βάθος το προφίλ αυτών των διαδικασιών και των ατόμων που εμπλέκονται σε αυτές, είμαστε σε θέση να εξάγουμε συμπεράσματα για αυτές και να τις βελτιώσουμε. Παρακάτω θα μελετήσουμε το μείζον ζήτημα του process mining, όπου μέσα από process model αναλύει διαδικασίες από τα αρχεία καταγραφής, με σκοπό να εντοπίσει παθολογίες και να τις αντιμετωπίσει.

1.2 ΑΝΤΙΚΕΙΜΕΝΟ ΤΗΣ ΔΙΠΛΩΜΑΤΙΚΗΣ

Αντικείμενο της διπλωματικής είναι η ανάλυση και αξιολόγηση των επιχειρησιακών διαδικασιών σε Οργανισμούς Τοπικής Αυτοδιοίκησης με τη χρήση τεχνικών εξόρυξης διαδικασιών, έτσι θα μπορούμε να εντοπίσουμε προβλήματα και να βελτιώσουμε τις διεργασίες. Τα δεδομένα της εργασίας προέρχονται από την Τεχνική Υπηρεσία πέντε Ολλανδικών δήμων και παρουσιάζονται σε μορφή event logs, με βασικό στόχο να αναλυθούν, να εξαχθούν και να αξιολογηθούν τα μοντέλα διαδικασιών.

Οι τεχνικές εξόρυξης διαδικασιών και μοντέλων διαδικασιών έγιναν με την χρήση του open source εργαλείου Prom Tools.

Με την διπλωματική στοχεύουμε να δώσουμε απαντήσεις σε ερωτήματα που αφορούν αυτούς τους δήμους αλλά και σε άλλους οργανισμούς, αποσκοπώντας στην βελτιστοποίηση της απόδοσης αυτών.

Κάποια από τα ερωτήματα αφορούν:

- Ποιές είναι οι πιο συνηθισμένες εργασίες στο κάθε δήμο αλλά και οι πιο σπάνιες?
- Ποιοι οργανισμοί/υπάλληλοι διεκπεραιώνουν τις περισσότερες εργασίες αλλά και τις πιο λίγες αντίστοιχα?
- Ποιοί δήμοι έχουν την υψηλότερη παραγωγικότητα και απόδοση?
- Πως αλληλεπιδρούν οι εργασίες και οι υπάλληλοι στους 5 δήμους μεταξύ τους?
- Κατά πόσο συμφωνούν τα μοντέλα διαδικασιών με τα αρχεία καταγραφής και τις μετρήσεις που έχουμε?

1.3 ΔΙΑΡΘΡΩΣΗ ΤΗΣ ΔΙΠΛΩΜΑΤΙΚΗΣ

Για την καλύτερη οργάνωση και δομή του σκελετού της διπλωματικής, στο κεφάλαιο 2 θα ξεκινήσουμε με το θεωρητικό υπόβαθρο της εξόρυξης διαδικασιών. Εκεί θα αναφέρουμε την πορεία από την πρώιμη επιχειρηματική ευφεία που αναπτύχθηκε για την αύξηση της απόδοσης των επιχειρήσεων, προς την ανάγκη για διαχείριση επιχειρηματικών διαδικασιών με ευέλικτο τρόπο. Με τεχνικές όπως η εξόρυξη δεδομένων και διαδικασιών με τα εργαλεία τους, όπως τα αρχεία καταγραφής συμβάντων και τα μοντέλα διαδικασιών, θα είμαστε σε θέση να μελετήσουμε και να διαχειριστούμε τις διαδικασίες που εκτελούνται και να τις βελτιώσουμε.

Στη συνέχεια το 3ο κεφάλαιο αποσκοπεί να μας εισάγει από την θεωρία στην πράξη, όπου στην αρχή του παρουσιάζει έναν γενικό οδηγό για την εξόρυξη διαδικασιών που μπορεί να εφαρμοστεί σε οποιαδήποτε κατάσταση ή επιχείρηση/οργανισμό βρεθούμε. Στη συνέχεια του 3ου κεφαλαίου στόχος είναι να συγκεκριμενοποιήσουμε τα βήματα και τις τεχνικές που ακολουθούμε στην πλατφόρμα Prom Tools. Ξεκινώντας στα απλά σκαλοπάτια βήμα βήμα από την εισαγωγή και ανάλυση των αρχείων καταγραφής συμβάντων, μέχρι τις τεχνικές για επεξεργασία και έλεγχο των δεδομένων. Ενώ με τα εργαλεία που διαθέτουμε μπορούμε να βρούμε κατά πόσο συμφωνούν η θεωρία με την πράξη, ελέγχοντας αν από τα αρχεία καταγραφής συμβάντων, παίρνουμε τα αναμενόμενα αποτελέσματα, στην αναπαράσταση των στοιχείων με τα μοντέλα διαδικασιών. Με αυτόν τον τρόπο συμπληρώνεται το θεωρητικό στοιχείο της διπλωματικής.

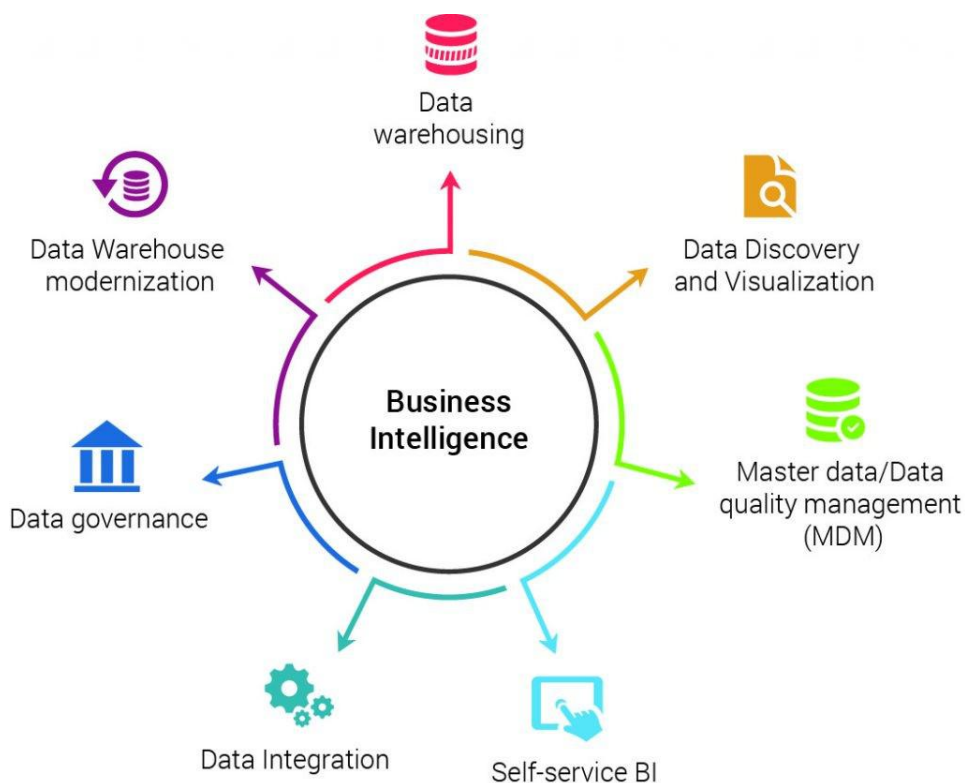
Με σκοπό στο 4ο κεφάλαιο να εισέλθουμε στο πειραματικό μέρος με τους Ολλανδικούς δήμους και την εξόρυξη διαδικασιών σε αυτούς. Εφαρμόζοντας τις τεχνικές που αναφέρθηκαν στα προηγούμενα κεφάλαια περνάμε βήμα βήμα από τα λόγια στα έργα για να μελετήσουμε τα δεδομένα, να τα επεξεργαστούμε κατάλληλα και να εξάγουμε συμπεράσματα. Στο 5 και τελευταίο κεφάλαιο, σχολιάζουμε τα συμπεράσματα και με βάση αυτά, απαντάμε στα καίρια ερωτήματα της διπλωματικής για τους δήμους, δείχνοντας την συνεισφορά και την ανάγκη της εξόρυξης διαδικασιών στο κλάδο των επιχειρήσεων.

ΚΕΦΑΛΑΙΟ 2

2.1 ΕΠΙΧΕΙΡΗΜΑΤΙΚΗ ΕΥΦΕΙΑ(BUSINESS INTELLIGENCE)

Όπως αναφέραμε και προηγουμένως κάθε οργανισμός και κάθε επιχείρηση έχει ως στόχο την βελτιστοποίηση των διαδικασιών που εκτελούνται σε αυτήν. Μία επιχείρηση, μέσα από την υψηλή παραγωγή και διάθεση υπηρεσιών στοχεύει στην αύξηση του κέρδους. Ενώ από την πλευρά του ένας οργανισμός είτε κρατικός είτε ιδιωτικός, αποσκοπεί στην ολοένα και καλύτερη εξυπηρέτηση των αναγκών που έχει κάθε άτομο και βελτίωση των υπηρεσιών που προσφέρει.

Για την επίτευξη των παραπάνω στόχων, τα τελευταία χρόνια έχει αναπτυχθεί ένας νέος όρος που ακούει στο όνομα επιχειρηματική ευφυΐα (BI). Αυτή η τεχνική έχει ως σκοπό να βοηθήσει μία επιχείρηση να αυξήσει τις δραστηριότητες που φέρει εις πέρας για να βελτιώσει την απόδοση, την λήψη αποφάσεων και την προγνωστική ανάλυση της. Αυτός ο νέος για την εποχή όρος συνδυάζει πεδία όπως ανάλυση επιχειρήσεων, εξόρυξη και εικονικοποίηση δεδομένων με τη βοήθεια εργαλείων δεδομένων και υποδομών, με σκοπό να οδηγήσει μία επιχείρηση, μικρή ή μεγάλη, να λαμβάνει σωστές αποφάσεις με βάση τα δεδομένα που έχει στην κατοχή της. Η επιχειρηματική ευφυΐα (BI) ξεκινάει να υπάρχει έντονα στον κλάδο των επιχειρήσεων από την δεκαετία του 1950 με την ψηφιακή επανάσταση της εποχής, δείχνοντας από τότε τον κομβικό ρόλο που θα έπαιζε για τις επιχειρήσεις. Τις επόμενες δεκαετίες αρκετές εταιρίες, που σήμερα είναι τεχνολογικοί κολοσσοί όπως η IBM και η SAP, υιοθέτησαν αυτήν την τεχνική ως επιχειρηματική στρατηγική τους για να υπερισχύσουν έναντι των ανταγωνιστών τους. Για να φτάσουμε στο σήμερα όπου εκατοντάδες εταιρίες ενστερνίζονται αυτήν την τεχνική. [3]

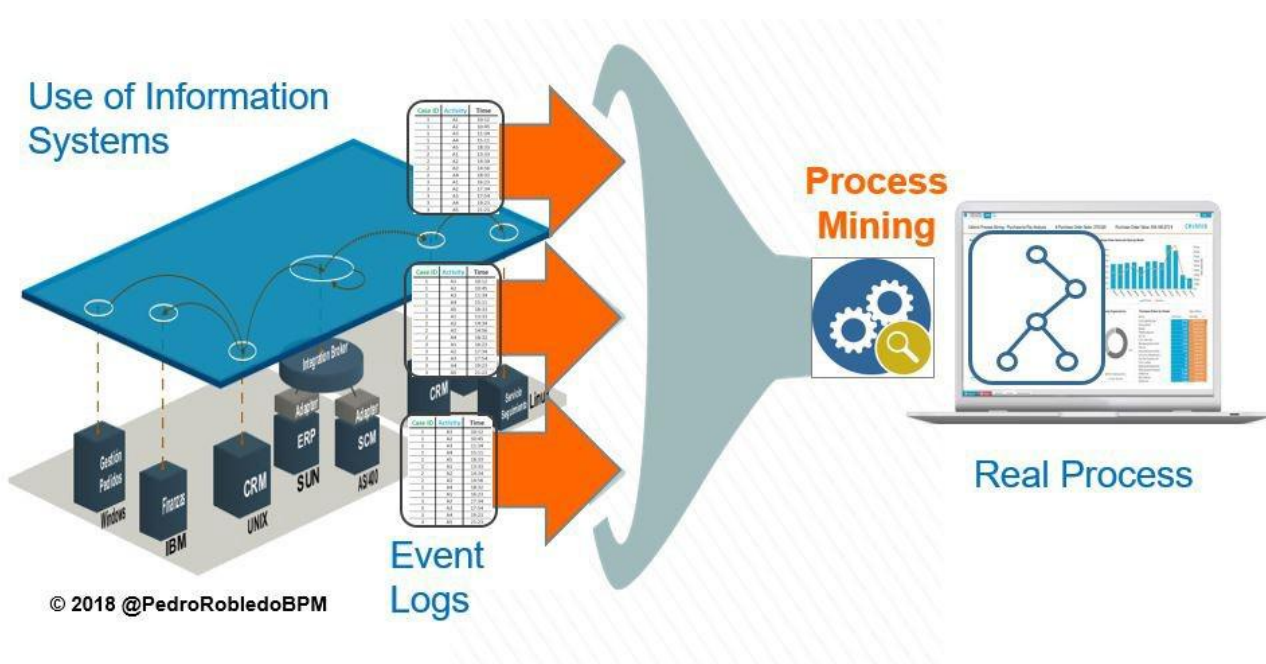


Εικόνα 1: Business intelligence in architecture and technology

Μερικά από τα ειδικά οφέλη που αποκομίζουν οι επιχειρήσεις και οργανισμοί κατά τη χρήση του ΒΙ περιλαμβάνουν:

- Αυξημένη αποτελεσματικότητα των επιχειρησιακών διαδικασιών.
- Πληροφορίες για τη συμπεριφορά των πελατών και τα μοτίβα αγορών.
- Ακριβής παρακολούθηση πωλήσεων, μάρκετινγκ και οικονομικών επιδόσεων.
- Διαγραφή κριτηρίων αναφοράς βάσει ιστορικών και τρεχόντων δεδομένων.
- Άμεσες ειδοποιήσεις σχετικά με ανωμαλίες δεδομένων και ζητήματα πελατών.
- Αναλύσεις που μπορούν να κοινοποιηθούν σε πραγματικό χρόνο σε διάφορα τμήματα.

Στο παρελθόν, τα εργαλεία επιχειρηματικής ευφυΐας χρησιμοποιούνταν κυρίως από αναλυτές δεδομένων και χρήστες πληροφορικής. Τώρα, οι πλατφόρμες ΒΙ αυτοεξυπηρέτησης καθιστούν την επιχειρηματική ευφυΐα διαθέσιμη σε όλους, από στελέχη έως ομάδες επιχειρήσεων. Σήμερα η επιχειρηματική ευφυΐα χρησιμοποιείται σε πολλούς τομείς όπως οι χρηματοοικονομικές υπηρεσίες, η υγεία, η παραγωγή και οι μετακινήσεις.



Εικόνα 2: Η πορεία του Process mining

2.2 ΔΙΑΧΕΙΡΙΣΗ ΕΠΙΧΕΙΡΗΜΑΤΙΚΩΝ ΔΙΑΔΙΚΑΣΙΩΝ(BPM)

Καθώς όμως η επιχειρηματική ευφυΐα χρησιμοποιείται κατά κόρον με την εξόρυξη και εικονοκοποίηση δεδομένων, την ανάλυση και σχεδίαση της ,πιο αποδοτικής, στρατηγικής τους και την αναζήτηση για νέες καλύτερες αποφάσεις οδηγούμενες από την ανάλυση στοιχείων που είχαν, παρατήρησαν ότι η επιχειρηματική ευφυΐα είχε κάποιους μείζων περιορισμούς,

Στην ουσία ότι γίνεται σε έναν οργανισμό μπορεί να θεωρηθεί ως μία διαδικασία. Η επιτυχία μίας επιχείρησης συνδέεται άμεσα με το πόσο καλά διαχειρίζεται τις επιχειρηματικές διαδικασίες της. Η αποτελεσματική διαχείριση απαιτεί κατανόηση της ποιότητας και της επικαιρότητας του τρόπου με τον οποίο αυτές οι διαδικασίες εκτελούνται. Ένας σημαντικός περιορισμός των εργαλείων της επιχειρηματικής ευφυΐας και ανακάλυψης δεδομένων είναι ότι οι μετρήσεις και αναλύσεις δεν έχουν σύνδεση με την επιχειρησιακό περιβάλλον, καθιστώντας δύσκολο για τον χρήστη να ερμηνεύσει και να δώσει νόημα σε αυτές τις μετρήσεις. Αυτά τα εμπόδια φανερώουν την ανάγκη για μία αυτοματοποιημένη ευφυΐα διαδικασιών(PI). Αυτός ο όρος σηματοδοτεί το επόμενο βήμα στην επιχειρηματική ευφυΐα και περιλαμβάνει την περισυλλογή δεδομένων με σκοπό την διαχείριση των επιχειρηματικών διαδικασιών(BPM) και ροών εργασίας για να αυξήσει την αποδοτικότητα των διαδικασιών. [4]

Κατά την διαχείριση επιχειρηματικών διαδικασιών έχουμε να κάνουμε με ένα κλάδο που χρησιμοποιεί διάφορες μεθόδους για την ανακάλυψη, τη μοντελοποίηση, την ανάλυση, τη μέτρηση, τη βελτίωση και τη βελτιστοποίηση των επιχειρηματικών διαδικασιών. Μια επιχειρηματική διαδικασία συντονίζει τη συμπεριφορά των ανθρώπων, των συστημάτων, των πληροφοριών και των πραγμάτων για την παραγωγή επιχειρηματικών αποτελεσμάτων για την υποστήριξη μιας επιχειρηματικής στρατηγικής. Οι διεργασίες μπορούν να είναι δομημένες και επαναλαμβανόμενες ή μη δομημένες και μεταβλητές. Αν και δεν απαιτείται, οι τεχνολογίες χρησιμοποιούνται συχνά με το BPM. Το BPM είναι το κλειδί για την ευθυγράμμιση των επενδύσεων IT/OT με την επιχειρηματική στρατηγική.

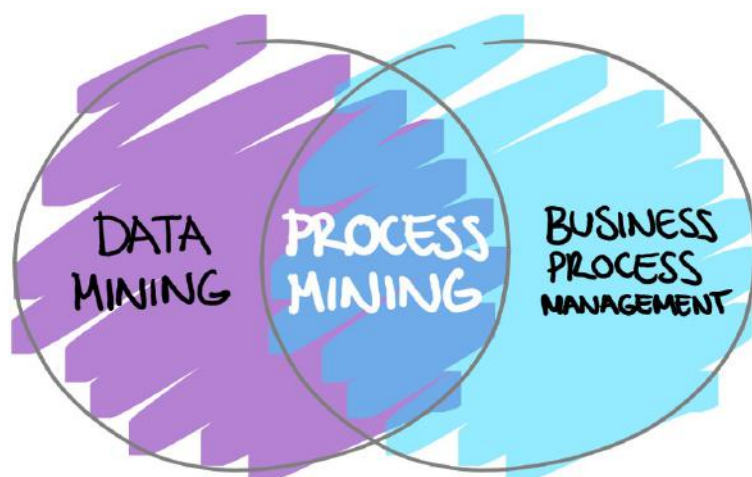
Ως προσέγγιση, το BPM βλέπει τις διαδικασίες ως σημαντικά περιουσιακά στοιχεία ενός οργανισμού που πρέπει να κατανοηθούν, να διαχειρίζονται και να αναπτυχθούν για να ανακοινωθούν και να παραδοθούν προϊόντα και υπηρεσίες προστιθέμενης αξίας σε πελάτες ή πελάτες. Αυτή η προσέγγιση μοιάζει πολύ με άλλες μεθοδολογίες διαχείρισης συνολικής ποιότητας ή διαδικασίας συνεχούς βελτίωσης. [5]

Με το στοιχείο της ευφυΐας που περιλαμβάνει αυτός ο όρος σηματοδοτεί την ανάγκη για ένα επιπλέον επίπεδο διαφάνειας, που πρέπει να προστεθεί, για να γίνει πλήρως κατανοητό το τι πρέπει να γίνει σε μία επιχείρηση για να διαφυλάξει την αξιοπιστία και ποιότητα των διαδικασιών της. Η ευφυΐα διαδικασιών(PI) θέτει ως κύριο στόχο να βελτιώσει τις υπηρεσίες και τα αγαθά, που προσφέρει μία επιχείρηση ,μέσα από την ουσιαστική κατανόηση και ανάλυση των διαδικασιών

της. Ουσιαστικά αυτός ο κλάδος ασχολείται με τις διαδικασίες (ενέργειες) μίας επιχείρησης, παρατηρώντας και αναλύοντας τες, για να βρει τι χρειάζεται, ώστε να γίνουν ακόμα πιο αποδοτικές.

2.3 ΕΞΟΡΥΞΗ ΔΕΔΟΜΕΝΩΝ

Ένας άλλος πολύ σημαντικός κλάδος που άρχισε να αναπτύσσεται σιγά σιγά ήταν η εξόρυξη δεδομένων, παρότι δεν τον γνώριζε κανείς μέχρι τη δεκαετία του 1990, γνώρισε ραγδαία αύξηση τα επόμενα χρόνια μέχρι σήμερα που έχει γιγαντωθεί. Λόγω των αναρίθμητων διαδικασιών που εκτελούνταν ακατάπαυστα, ο τομέας της εξόρυξης δεδομένων είχε έρθει για να μελετήσει τα δεδομένα που υπήρχαν από τις διαδικασίες. Η εξόρυξη δεδομένων είναι η διαδικασία εύρεσης ανωμαλιών, προτύπων και συσχετίσεων μέσα σε μεγάλα σύνολα δεδομένων για την πρόβλεψη των αποτελεσμάτων. Χρησιμοποιώντας ένα ευρύ φάσμα τεχνικών, μπορείτε να χρησιμοποιήσετε αυτές τις πληροφορίες για να αυξήσετε τα έσοδα, να μειώσετε το κόστος, να βελτιώσετε τις σχέσεις με τους πελάτες, να μειώσετε τους κινδύνους και πολλά άλλα. Μέσα σε αυτήν την τεχνολογία, η εξόρυξη συνδυάζεται και με άλλες επιστήμες για να μπορεί να βελτιώσει τα αποτελέσματα που παράγει όπως στατιστική (η αριθμητική μελέτη των σχέσεων δεδομένων), τεχνητή νοημοσύνη (ανθρώπινη νοημοσύνη που εμφανίζεται από λογισμικό ή/και μηχανές) και μηχανική μάθηση (αλγόριθμοι που μπορούν να μάθουν από δεδομένα για να κάνουν προβλέψεις). Όσο περνάνε τα χρόνια η τεχνολογία εξόρυξης δεδομένων συνεχίζει να εξελίσσεται για να συμβαδίζει με τις απεριόριστες δυνατότητες των μεγάλων δεδομένων και την προσιτή υπολογιστική ισχύ. [6]



Εικόνα 3: Σύνδεση ανάμεσα σε εξόρυξη δεδομένων και διαχείριση επιχειρηματικών διαδικασιών

2.4 ΕΞΟΡΥΞΗ ΔΙΑΔΙΚΑΣΙΩΝ

Την σύνδεση ανάμεσα στην εξόρυξη δεδομένων (data mining) και την διαχείριση επιχειρηματικών διαδικασιών(PI) έρχεται να κάνει η εξόρυξη διαδικασιών(PM) ,η οποία αποτελεί και το κύριο θέμα αυτής της διπλωματικής.Αυτός ο νεοφυής κλάδος πηγάζει από την ευφυΐα διαδικασιών που αναφέραμε παραπάνω και έχει ως σκοπό,την ανάλυση,παρακολούθηση και βελτίωση σε βάθος αληθινών διαδικασιών,που λαμβάνουν χώρα καθημερινά σε εκατομμύρια επιχειρήσεις σε όλο το κόσμο.Θα μπορούσαμε να πούμε ότι η κύρια διαφορά ανάμεσα στον “πατέρα” αυτής της τεχνικής , διαχείριση επιχειρηματικών διαδικασιών ,που ξεκίνησε αυτόν τον κλάδο και το “παιδί” αυτού,την εξόρυξη διαδικασιών(PM) είναι ότι το δεύτερο ασχολείται με τα βαθύτερα αίτια που οδήγησαν σε μία κατάσταση,ενώ το πρώτο με τις τεχνικές που περιλαμβάνει,μπορεί απλά να διαπιστώσει το πρόβλημα χωρίς να είναι σε θέση να το διορθώσει.

2.4.1 ΕΙΣΑΓΩΓΗ

Ο κλάδος της εξόρυξης διεργασιών εμφανίστηκε στα τέλη της δεκαετίας του 1990 στο Τεχνολογικό Πανεπιστήμιο του Αϊντχόβεν με το πρωτοποριακό έργο του prof.dr.ir. Ο Γουίλ βαν ντερ Άαλστ. Με τον καιρό η πειθαρχία ωρίμασε και σήμερα, υπάρχουν πάνω από 35 εμπορικοί προμηθευτές εξόρυξης διεργασιών και χιλιάδες οργανισμοί που εφαρμόζουν με επιτυχία την εξόρυξη διεργασιών. Η Task Force IEEE για την εξόρυξη διεργασιών και το πλαίσιο ανοιχτού κώδικα ProM διαδραμάτισαν βασικό ρόλο στην ανάπτυξη του κλάδου. [7]

Πιο αναλυτικά τα γεγονότα που ακολούθησαν ,με χρονολογική σειρά για να γίνει η εξόρυξη διεργασιών αυτό είναι σήμερα,είναι τα εξής:

ΧΡΟΝΟΛΟΓΙΟ ΕΞΟΡΥΞΗΣ ΔΙΕΡΓΑΣΙΩΝ

- 1999 Η αρχή της αναζήτησης για την εξόρυξη διεργασιών στο πανεπιστήμιο του Αϊντχόβεν
- 2000 Δημιουργία του αλγόριθμου alpha miner για εξαγωγή μοντέλου διαδικασιών
- 2001 Δημιουργία του αλγόριθμου heuristic miner
- 2004 Βασισμένος σε στοιχεία έλεγχος συμμόρφωσης
- 2005 Εξόρυξη αποφάσεων
- 2006 Εξόρυξη οργανισμών
- 2007 Πρώτη εταιρία για εξόρυξη διεργασιών (Futura Pi)

- 2009 Πρώτες ομάδες εργασίας στο τομέα της εξόρυξης διαδικασιών και δημιουργία Fluxicon πλατφόρμα εξόρυξης διεργασιών
- 2010 Ευθυγραμμισμένη έλεγχος συμμόρφωσης
- 2011 Δημιουργία πρώτου βιβλίου εξόρυξης διεργασιών και δημιουργία Celonis πλατφόρμα εξόρυξης διεργασιών
- 2014 Εξόρυξη διεργασιών της Coursera MOOC
- 2016 Κυκλοφορία βιβλίου “Εξόρυξη διεργασιών:Η Επιστήμη δεδομένων σε δράση”
- 2018 Πάνω από 30 εταιρίες με ειδίκευση την εξόρυξη διεργασιών και τα εργαλεία της.Το λογισμικό Celonis αποτελεί την ναυαρχίδα στο τομέα.
- 2019 ICPM 2019 :Πρώτο συνέδριο εξόρυξης διεργασιών

2.4.2 ΙΔΙΟΤΗΤΕΣ ΚΑΙ ΕΦΑΡΜΟΓΕΣ

Η εξόρυξη διαδικασιών ως τομέας είναι σε θέση να απαντήσει σε 4 μείζων ερωτήματα

- >τι πραγματικά συνέβη σε μία διαδικασία(εργασία)
- >αποκλίσεις από την αρχική της σχεδίαση
- >σε ποία σημεία υπάρχει συμφόρηση(bottleneck)
- >βελτίωση της διαδικασίας

Η εξόρυξη διαδικασιών χρησιμοποιείται σε μεγάλο βαθμό,σε μία ευρεία γκάμα πεδίων.Την μεγαλύτερη συνεισφορά την έχει σε σημαντικούς τομείς όπως οι:

Επιχειρήσεις:όπως είναι και το κύριο θέμα της διπλωματικής,το process mining εφαρμόζεται αρκετά στον επιχειρηματικό τομέα.Οι τεχνικές που διαθέτει βελτιστοποιούν τις διεργασίες που λαμβάνουν χώρα σε μία επιχείρηση.Οι αλγόριθμοι βοηθάνε στην καλύτερη ανάλυση και μελέτη των εργασιών,ενώ με αυτόν τον τρόπο μπορεί να αξιολογηθεί η απόδοση και τα πιθανά σημεία που την επιβραδύνουν.Τα μοντέλα διαδικασιών προσφέρουν μία οπτική παρουσίαση των διεργασιών,εξάγοντας ένα πρώτο συμπέρασμα για το βαθμό που συμφωνεί με το αρχείο καταγραφών.Τα τελευταία χρόνια οι κολοσσοί και αργότερα όλες οι επιχειρήσεις εγκαθιδρύουν τέτοιους τομείς,με σκοπό να αξιολογηθεί ο τρόπος λειτουργίας.

Στον τομέα των επιχειρήσεων και των οργανισμών η εξόρυξη διαδικασιών μπορεί να προσφέρει οφέλη όπως:

→ Δυνατότητα εξοικονόμησης, βλέποντας ακριβώς πού είναι η σπατάλη και η απώλεια που γίνεται σε μία επιχείρηση

→ Πλήρη γνώση των διεργασιών που λαμβάνουν χώρα, πότε ξεκινάνε και πότε τελειώνουν, ποία θα είναι η πορεία τους και αν υπάρχει κάποιο πρόβλημα σε αυτές.

→ Ανακάλυψη αδύνατων σημείων και ευπαθειών του συστήματος μας, που μπορεί να δημιουργήσουν προβλήματα

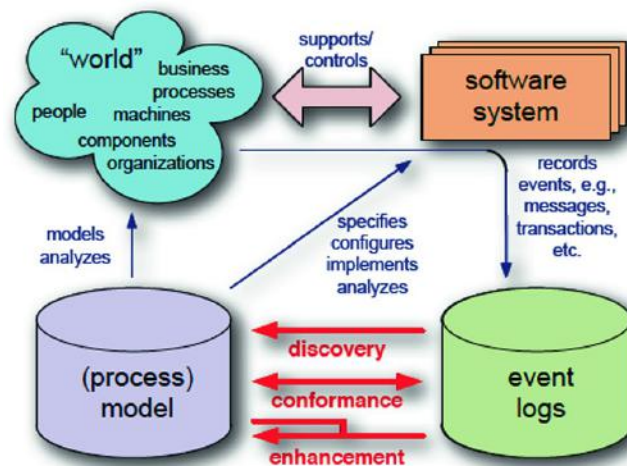
→ Παρατήρηση του τρόπου που συνδέονται οι διαδικασίες και πως αλληλεπιδρούν με τους καταναλωτές, δημιουργώντας μοτίβα για ανάλυση

Υγεία: τα τελευταία χρόνια ,με το διαδίκτυο μπορούμε να προσφέρουμε ιατρική περίθαλψη σε άτομα με ανάγκες, που παλαιότερα αυτό ήταν αρκετά δύσκολο. Για παράδειγμα το ιστορικό του ασθενούς βρίσκεται ψηφιακά, η συνταγογράφηση γίνεται ηλεκτρονικά, ακόμα και η παρακολούθηση και η περίθαλψη μπορεί να γίνει ηλεκτρονικά, ειδικά τους τελευταίους μήνες με την πανδημία, γινόταν προσπάθειες να γίνονται όλα ηλεκτρονικά. Με την επικρατούσα κατάσταση καθημερινά γίνονται εκατομμύρια τέτοιες διαδικασίες. Εφόσον έχουμε πολλές διαδικασίες που εκτελούνται ,θα πρέπει να γίνετε αξιολόγηση αυτών, αναζήτηση για πιθανά προβλήματα και βελτίωση των υπηρεσιών που διατίθενται. Για αυτούς τους λόγους η εξόρυξη διαδικασιών στο τομέα της υγείας είναι απαραίτητη και θα πρέπει όλοι οι οργανισμοί/επιχειρήσεις να εφαρμόζουν αυτές τις τεχνικές.

Ανθρώπινη δραστηριότητα: όλοι οι άνθρωποι καθημερινά σε όλο το κόσμο εκτελούν πολλές δραστηριότητες , από τις πιο απλές μέχρι τις πιο σύνθετες. Ο καθένας έχει τις δικές του δραστηριότητες που επαλαμβάνονται καθημερινά, ενώ πολλοί άνθρωποι αναζητούν τον τρόπο , να γίνονται οι δραστηριότητες πιο αποδοτικά και σε λιγότερο χρόνο. Ενδιαφέρον για αυτό το ζήτημα υπάρχει από πολλούς ειδικούς, οι οποίοι εφαρμόζοντας τεχνικές εξόρυξης διεργασιών, είναι σε θέση να αναλύσουν δραστηριότητες και συμπεριφορές και να τις βελτιώσουν, συμβάλλοντας στην ευημερία των πολιτών. Αυτό συνολικά θα επηρεάσει και όλη την κοινωνία ,βοηθώντας στην συνολική εξέλιξη της και την ανάπτυξη του βιοτικού επιπέδου της. Αξίζει βέβαια να σημειωθεί ότι το process mining σε δραστηριότητες ανθρώπων, εξυπηρετεί και σκοπούς διαφήμισης των επιχειρήσεων, με στόχο να προσφέρουν εξατομικευμένα προϊόντα και υπηρεσίες στους πολίτες.

Μεταφορές: κάθε μέρα οι άνθρωποι χρησιμοποιούν μέσα για να μεταφερθούν από ένα σημείο σε ένα άλλο. Αυτό γίνεται είτε με δικό τους μέσο είτε μέσα από συγκοινωνίες πληρώνοντας για την μεταφορά. Εκατομμύρια τέτοιες διαδικασίες συμβαίνουν καθημερινά σε πολλές διαφορετικές περιοχές, σε διάφορες ώρες και με διαφορετικά μέσα. Σαν κλάδος είναι γιγαντιαίος και συναντάτε καθημερινά στη ζωή πολλών ανθρώπων. Είναι τεράστιος ο όγκος δεδομένων που συλλέγονται από

αυτήν την δραστηριότητα και με πολλές διαφοροποιήσεις, ανάλογα το μέσο και τα χαρακτηριστικά που διαθέτει. Η εξόρυξη διαδικασιών σαν τομέας μπορεί να φιλτράρει αυτά τα στοιχεία, να βρει επαναλαμβανόμενα μοτίβα, παθολογίες και άλλα ζητήματα που προκύπτουν. Με τις τεχνικές που διαθέτει μπορεί να δείξει στους αρμόδιους τα ζητήματα και να βοηθήσει στην επίλυση τους, προσφέροντας έτσι καλύτερες υπηρεσίες για τους ενδιαφερόμενους.



Εικόνα 4: Λειτουργία του process mining στην καθημερινότητα

Η εφαρμογή του Process Mining σε έναν οργανισμό προσφέρει τις ακόλουθες δυνατότητες: [8]

-Αυτοματοποιημένη ανακάλυψη μοντέλων διεργασιών, εξαιρέσεων και περιπτώσεων διεργασιών(περιπτώσεων) μαζί με βασικές συχνότητες και στατιστικά στοιχεία.

-Αυτοματοποιημένη ανακάλυψη και ανάλυση των αλληλεπιδράσεων με τους πελάτες, καθώς και ευθυγράμμιση με εσωτερικές διαδικασίες.

-Παρακολούθηση βασικών δεικτών απόδοσης χρησιμοποιώντας πίνακες εργαλείων σε πραγματικό χρόνο.

-Βελτίωση υφιστάμενων ή προηγούμενων μοντέλων διεργασιών με χρήση πρόσθετων δεδομένων από αποθηκευμένες εγγραφές.

-Συνδυασμός διαφορετικών μοντέλων διεργασιών που αλληλεπιδρούν μεταξύ τους σε ένα ενιαίοπάνελεξόρυξης διεργασιών.

-Υποστήριξη για την οπτικοποίηση του τρόπου με τον οποίο οι διαδικασίες συνεισφέρουν στην επιχειρηματική αξία (όπως τα επιχειρηματικά μοντέλα λειτουργίας) - πλαισίωση των διαδικασιών.

-Αποτελεσματική συνεργασία μεταξύ Επιχειρήσεων και Πληροφορικής.

-Τυποποίηση επιχειρηματικών διαδικασιών.

-Βελτίωση της λειτουργικής αριστείας με τη βελτιστοποίηση των διαδικασιών.

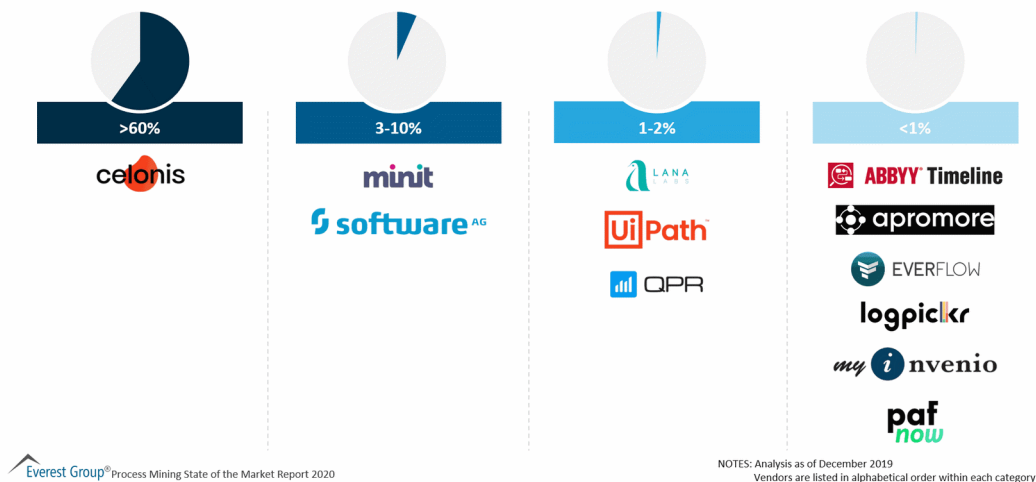
2.4.3 ΕΡΓΑΛΕΙΑ ΕΞΟΡΥΞΗΣ ΔΙΑΔΙΚΑΣΙΩΝ



Εικόνα 5: Κύριοι δημιουργοί λογισμικών εξόρυξης διαδικασιών ανά τον κόσμο

Για τη εξόρυξη διαδικασιών (process mining) μπορούμε να χρησιμοποιήσουμε διάφορα εργαλεία και λογισμικά, τόσο open source όσο και με πληρωμή, τα οποία μπορεί να διαθέτουν περισσότερα χαρακτηριστικά. Παραπάνω βλέπουμε διάφορες επιλογές, ακόμα και το Prom Tools που χρησιμοποιούμε στα παραδείγματα μας, αλλά και εναλλακτικές αν προτιμάμε κάτι άλλο. Η Ευρώπη αντιπροσωπεύει το μεγαλύτερο μέγεθος αγοράς λόγω της ευρείας αποδοχής των καινοτομιών και της επερχόμενης τεχνολογίας αναλυτικών στοιχείων μεταξύ του μεγάλου αριθμού προμηθευτών λογισμικού Process Mining που δραστηριοποιούνται στην περιοχή. Η Γερμανία και η Ολλανδία είναι οι κορυφαίες χώρες που συνεισφέρουν στην αγορά ανάλυσης διεργασιών στην Ευρώπη. Η Βόρεια Αμερική παρουσιάζει τον υψηλότερο ρυθμό ανάπτυξης, αντιπροσωπεύοντας τη μεγαλύτερη ευκαιρία στην αγορά ανάλυσης διεργασιών. [9]

Process Mining Software Vendors' Market Share by License Revenue



Εικόνα 6: Μερίδιο αγοράς προμηθευτών λογισμικού εξόρυξης διεργασιών

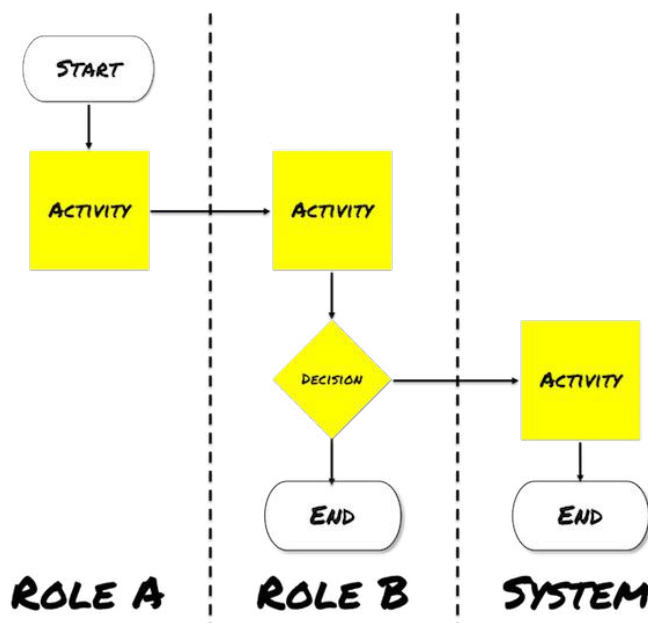
Αξίζει τέλος να σημειωθεί ότι τα τελευταία χρόνια όλο και περισσότερα κράτη ξεκινάνε να επενδύουν σε αυτόν τον τομέα. Οι προοπτικές που προσφέρει είναι τεράστιες και τα οφέλη ουκ ολίγα για τις επιχειρήσεις και τα κράτη συνεπώς. Στην Ελλάδα έχουν αρχίσει προσπάθειες για εφαρμογή τέτοιων τεχνικών και στο δημόσιο τομέα, με σκοπό της βελτίωσης του τρόπου εκτέλεσης εργασιών. Πρόσφατα έχει αυξηθεί η προσπάθεια για ανάπτυξη του κλάδου, μέσα από σημαντική χρηματοδότηση, κάνοντας την αγορά των process analytics αρκετά ελκυστική. Όπως είναι λογικό πολλές νέες θέσεις για data analyst, data scientist και data engineer έχουν δημιουργηθεί συγκεντρώνοντας το ενδιαφέρον από όλους τους κλάδους. Σαν επιστήμη συνδυάζει πολλά πεδία όπως πληροφορική, προγραμματισμό, οικονομικά και μπορεί να εφαρμοστεί σε οποιαδήποτε τομέα που προαναφέραμε, για παράδειγμα στην καθημερινή ζωή μας, την εργασία μας, τις μετακινήσεις μας ακόμα τον τρόπο ψυχαγωγίας που μας ενδιαφέρει. Η εξόρυξη διεργασιών συχνά παρερμηνεύεται ως πεδίο που σχετίζεται με την επιστήμη των δεδομένων. Η εξόρυξη διεργασιών θα πρέπει να θεωρείται ως γέφυρα μεταξύ της επιστήμης δεδομένων και της επιστήμης διεργασιών. Μέχρι το έτος 2018, σχεδόν 30+ εμπορικά διαθέσιμα εργαλεία εξόρυξης διεργασιών ήταν στην εικόνα. Το έτος 2019 ορίστηκε το πρώτο συνέδριο εξόρυξης διεργασιών. Σήμερα έχουμε πάνω από 35 προμηθευτές που προσφέρουν εργαλεία και τεχνικές για την ανακάλυψη διαδικασιών και τον έλεγχο συμμόρφωσης. Παρά το νεαρό της ηλικίας αυτού του κλάδου, είναι ο πιο ταχεία αναπτυσσόμενος στον τομέα της πληροφορικής και προσεγγίζει ενδιαφερόμενους από όλο το κόσμο, να ασχοληθούν μαζί του.



Εικόνα 7: Η αγορά της ανάλυσης διεργασιών

2.5 PROCESS MODEL KAI MINING ALGORITHMS

Η εξόρυξη διαδικασιών για να πετύχει τους στόχους της υλοποιεί ένα μοντέλο διαδικασιών (process model), το οποίο παρουσιάζει την αναμενόμενη λειτουργία μίας διαδικασίας και τι περιμένουμε να έχουμε ως τελικό αποτέλεσμα από αυτήν. Πιο συγκεκριμένα το μοντέλο διαδικασιών είναι η γραφική αναπαράσταση επιχειρηματικών διαδικασιών ή ροών εργασίας. Όπως ένα διάγραμμα ροής, τα μεμονωμένα βήματα της διαδικασίας καταρτίζονται, ώστε να υπάρχει μια συνολική επισκόπηση των εργασιών στη διαδικασία στο πλαίσιο του επιχειρηματικού περιβάλλοντος. Τα μοντέλα διαδικασιών προσφέρουν μία οπτικοποιημένη πληροφορία των διαδικασιών που εκτελούνται, ενώ έπειτα μέσα από ελέγχους μπορούμε να δούμε κατά πόσο συμφωνούν με τα αρχεία καταγραφής που διαθέτουμε. Έτσι με μία ματιά, βλέποντας το μπορούμε να ελέγξουμε αν το πραγματικό τελικό στάδιο μίας διεργασίας είναι το ίδιο με αυτό που προέβλεπε το μοντέλο διαδικασιών.



Εικόνα 8: Δομή διαγράμματος ροής ή μοντέλου διαδικασιών

Στο τομέα του process mining όπως αναφέραμε και προηγουμένως σημαντικό ρόλο παίζει και το process model ή αλλιώς μοντέλο διαδικασιών. Αποτελεί την οπτική αναπαράσταση του διαγράμματος διεργασιών ή της ροής ενεργειών. Για την παραγωγή αυτού του μοντέλου στον κόσμο της εξόρυξης διεργασιών, πρέπει να διαθέτουμε ειδικούς αλγόριθμους, όπου με την χρήση αυτών θα παράγεται το process tree. Συνήθως στις δοκιμές χρησιμοποιούμε τέσσερις τέτοιους αλγόριθμους, με σκοπό να καλύψουμε όλες τις διαφορετικές απαιτήσεις που έχουμε σε κάθε δοκιμή. Συγκεκριμένα τις περισσότερες φορές δουλεύουμε με τον Alpha miner, Heuristic miner, Inductive Visual miner και Fuzzy miner. Κάθε ένας από αυτούς έχει διαφορετικά χαρακτηριστικά και προσφέρει δυνατότητες, όπου ίσως οι υπόλοιποι δεν προσφέρουν. Σε κάθε περίπτωση πρέπει να βλέπουμε το πρόβλημα που αντιμετωπίζουμε και να διαλέγουμε τον βέλτιστο αλγόριθμο για αυτήν την δουλειά.

Παραπάνω παρουσιάστηκαν οι 4 αλγόριθμοι που χρησιμοποιούμε στο Prom Tools. Ας αναφέρουμε κάποια στοιχεία για αυτούς, ο alpha miner ήταν ο πρώτος αλγόριθμος που υλοποιήθηκε για αυτήν την χρήση και συγκεκριμένα το 2000. Ο α-αλγόριθμος ή α-miner είναι ένας αλγόριθμος που χρησιμοποιείται στην εξόρυξη διεργασιών, με στόχο την ανακατασκευή της αιτιότητας από ένα σύνολο γεγονότων που διαθέτουμε. Υλοποιήθηκε από τους van der Aalst, Weijters και Mārušter. Ο στόχος του Alpha miner είναι να μετατρέψει το log file σε ένα δίκτυο ροής εργασίας με βάση τις σχέσεις μεταξύ των διαφόρων δραστηριοτήτων στο αρχείο καταγραφής συμβάντων. Ένα αρχείο καταγραφής συμβάντων είναι ένα πολλαπλό σύνολο ιχνών (cases) και ένα ίχνος είναι μια ακολουθία ονομάτων δραστηριότητας (events). Έκτοτε έχουν παρουσιαστεί αρκετές επεκτάσεις ή τροποποιήσεις του, οι οποίες θα παρατίθενται παρακάτω. Πολλές από τις τεχνικές εξόρυξης διεργασιών έχουν επηρεαστεί από τον αρχικό α miner.

Μία από αυτές είναι ο Heuristic miner που υλοποιήθηκε λίγο αργότερα περίπου το 2001. Αποτελεί μία βελτιωμένη έκδοση του αρχικού αλγόριθμου και ήταν ο δεύτερος αλγόριθμος εξόρυξης διεργασιών, ακολουθώντας στενά τον αλγόριθμο άλφα. Αναπτύχθηκε από τον Δρ. Ton Weijters, ο οποίος χρησιμοποίησε μια ευρετική προσέγγιση για να αντιμετωπίσει πολλά προβλήματα με τον αλγόριθμο άλφα, καθιστώντας αυτόν τον αλγόριθμο πολύ πιο κατάλληλο στην πράξη. Ένα από τα πλεονεκτήματα είναι ότι ένα ευρετικό δίκτυο μπορεί να μετατραπεί σε άλλους τύπους μοντέλων διεργασιών, όπως ένα δίκτυο Petri για περαιτέρω ανάλυση στο ProM. Είναι καλό να τον προτιμάμε όταν έχουμε πραγματικά δεδομένα με όχι πάρα πολλά διαφορετικά συμβάντα ή όταν χρειάζεστε ένα μοντέλο Petri net για περαιτέρω ανάλυση στο ProM.

Ένας ακόμα αλγόριθμος που χρησιμοποιείται ευρέως στο Prom είναι ο Inductive Visual miner algorithm, ο οποίος αποτελεί μία αρκετά βελτιωμένη έκδοση των δύο προηγούμενων αλγορίθμων. Μας δίνει την δυνατότητα ζωντανής παρακολούθησης του μοντέλου διαδικασιών, βλέποντας σε πραγματικό χρόνο την ροή εκτέλεσης εργασιών. Ο αλγόριθμος αυτός πρόκειται για τον πιο αποδοτικό, στις περισσότερες περιπτώσεις, προσφέροντας αρκετές δυνατότητες παραμετροποίησης και καλύτερη παρουσίαση του τελικού process model. Στις μετρήσεις μας έχει χρησιμοποιηθεί αρκετά λόγω της ευκολίας στο χειρισμό και της παραμετροποίησης που προσφέρει σε ζωντανό χρόνο.

Τέλος έχουμε τον Fuzzy miner όπου εδώ έχουμε τον πιο καινούργιο αλγόριθμο στο είδος. Ο Fuzzy miner είναι ένας από τους νεότερους αλγόριθμους ανακάλυψης διεργασιών και αναπτύχθηκε από τον συνιδρυτή της Fluxicon Christian W. Gnther το 2007. Είναι ο πρώτος αλγόριθμος που αντιμετωπίζει άμεσα τα προβλήματα μεγάλου αριθμού δραστηριοτήτων και εξαιρετικά αδόμετης συμπεριφοράς. Ο Fuzzy miner χρησιμοποιεί μετρήσεις σημασίας/συσχέτισης για να απλοποιήσει διαδραστικά το μοντέλο διαδικασίας στο επιθυμητό επίπεδο αφαίρεσης. Σε σύγκριση με τον Heuristic miner, μπορεί επίσης να αφήσει εκτός λιγότερο σημαντικές δραστηριότητες (ή να τις κρύψει σε συμπλέγματα) εάν έχετε εκατοντάδες από αυτές. Προτιμάται σε περιπτώσεις όπου έχουμε πολύπλοκα και μη δομημένα δεδομένα καταγραφής ή όταν θέλουμε να απλοποιήσουμε το μοντέλο με διαδραστικό τρόπο. [10]

	Alpha Algorithm	Heuristics Miner	Genetic Miner	Fuzzy Miner
Integrity		✓	✓	
Uniqueness	✓	✓	✓	✓
Exceptions		✓	✓	
Real-Time	✓			
Model Integrity	✓	✓	✓	✓
Abstraction level				✓
Structured model				✓

Εικόνα 9: Σύγκριση μεταξύ των mining αλγορίθμων

2.6 ΑΡΧΕΙΑ ΚΑΤΑΓΡΑΦΗΣ ΣΥΜΒΑΝΤΩΝ

Όταν εκτελείται μία διεργασία στην επιχείρηση η διαδικασία αυτή, αφού ολοκληρωθεί αναγράφεται με τα βασικά στοιχεία της ,σε συγκεκριμένα αρχεία που περιλαμβάνουν όλες τις διεργασίες μίας επιχείρησης.Τα αρχεία αυτά που θα τα συναντήσουμε και παρακάτω ονομάζονται αρχείο καταγραφής συμβάντων(event logs).Είναι ένας βασικός πόρος στα σύγχρονα συστήματα που βοηθά στην παροχή πληροφοριών σχετικά με διεργασίες που λαμβάνουν χώρα και περιλαμβάνουν στοιχεία για κάθε διεργασία.Όπως για παράδειγμα το όνομα της δραστηριότητας,το κωδικό της(κάθε ενέργεια έχει διαφορετικό κωδικό και δεν μπορούν δύο διαδικασίες να έχουν τον ίδιο),τον χρόνο που χρειάστηκε για να ολοκληρωθεί και διάφορες άλλες ιδιότητες που ποικίλλουν ανάλογα με την ενέργεια.Κάποιες άλλες ενέργειες μπορεί να είναι το άτομο ή ο οργανισμός που εκτέλεσε την διεργασία,αν ολοκληρώθηκε και πόσο κόστισε.Μπορούν να υπάρχουν και άλλες ιδιότητες με σκοπό να κάνουν τις εργασίες συγκρίσιμες μεταξύ τους,για να βγαίνουν ευκολότερα συμπεράσματα. Αυτά τα αρχεία αποθηκεύουν τεράστιο αριθμό δεδομένων τα οποία είναι ιδιαίτερα χρήσιμα σε προσπάθειες ανάκτησης δεδομένων από ειδικούς και σε περιπτώσεις ανάλυσης και παρακολούθησης των ενεργειών μίας επιχείρησης,με σκοπό να διαπιστωθεί η απόδοση και οι αλλαγές που θα την ωφελούσαν.

2.6.1 ΔΟΜΗ ΑΡΧΕΙΩΝ

SessionID	Page	Timestamp	CookieID	DataCenter	SiteVersion
487434	portal.aspx	2016-01-01 15:34:01	A	phoenix	1.12
487434	dashboard.aspx	2016-01-01 15:34:15	A	phoenix	1.12
487434	purchaseorderreport.aspx	2016-01-01 15:34:30	A	phoenix	1.12
487435	portal.aspx	2016-01-01 14:01:10	B	phoenix	2
487435	help.aspx	2016-01-01 14:03:23	B	phoenix	2
487435	contactus.aspx	2016-01-01 14:04:07	B	phoenix	2
487436	portal.aspx	2016-01-01 17:11:17	A	phoenix	1.12
487436	myteam.aspx	2016-01-01 17:12:41	A	phoenix	1.12
487436	expensereports.aspx	2016-01-01 17:12:55	A	phoenix	1.12

Εικόνα 10: Παράδειγμα από αρχείο καταγραφής συμβάντων με ίχνη και διαδικασίες,όπου το καθένα περιλαμβάνει attributes

Όσον αφορά την δομή αυτών των αρχείων πρέπει να αναφέρουμε ότι δομούνται με την λογική ότι κάθε φορά που μία διεργασία τρέχει ,εμείς έχουμε ένα ίχνος(trace) αυτής.Κάθε ίχνος σε ένα log file μπορεί να αποτελείται από μία ακολουθία δραστηριοτήτων που εκτελούνται η μία μετά την άλλη.Μία ακολουθία μπορεί να διαφοροποιηθεί από μία άλλη με τις ίδιες διεργασίες,στην περίπτωση που οι ενέργειες τρέχουν σε διαφορετικούς χρόνους και με διαφορετική σειρά. Συμπεραίνουμε δηλαδή ότι σε τέτοιες περιπτώσεις με πάρα πολλές αναφορές ,παίζει τεράστιο ρόλο η σειρά με την οποία έγιναν οι ενέργειες και ποία ενέργεια διαδέχθηκε κάποια άλλη.

Σε γενικές γραμμές κάθε event log αποτελείται από 3 κύρια στοιχεία:

→ Αναγνωριστικό περίπτωσης(case id): κάθε συμβάν θα πρέπει να αναφέρεται σε μια περίπτωση (δηλαδή, παράδειγμα διαδικασίας).Χρησιμοποιείται για τον διαχωρισμό των υποθέσεων και για να μπορούμε με ευκολία να μελετήσουμε μία υπόθεση.

→ Δραστηριότητα (activity) :κάθε γεγονός πρέπει να σχετίζεται με μια δραστηριότητα. Τα γεγονότα αναφέρονται σε περιπτώσεις δραστηριοτήτων,δηλαδή εμφανίσεις δραστηριοτήτων στο αντίστοιχο μοντέλο διαδικασίας.

→ Χρονοσφραγίδα (timestamp) : τα γεγονότα εντός μιας υπόθεσης πρέπει να διατάσσονται σε χρονική σειρά εκτέλεσης.Επιπλέον,χρονοσφραγίδες δεν χρειάζονται μόνο για τη χρονική σειρά: είναι επίσης ζωτικής σημασίας για τη μέτρηση της διεργασίας και σύγκριση της με άλλες.

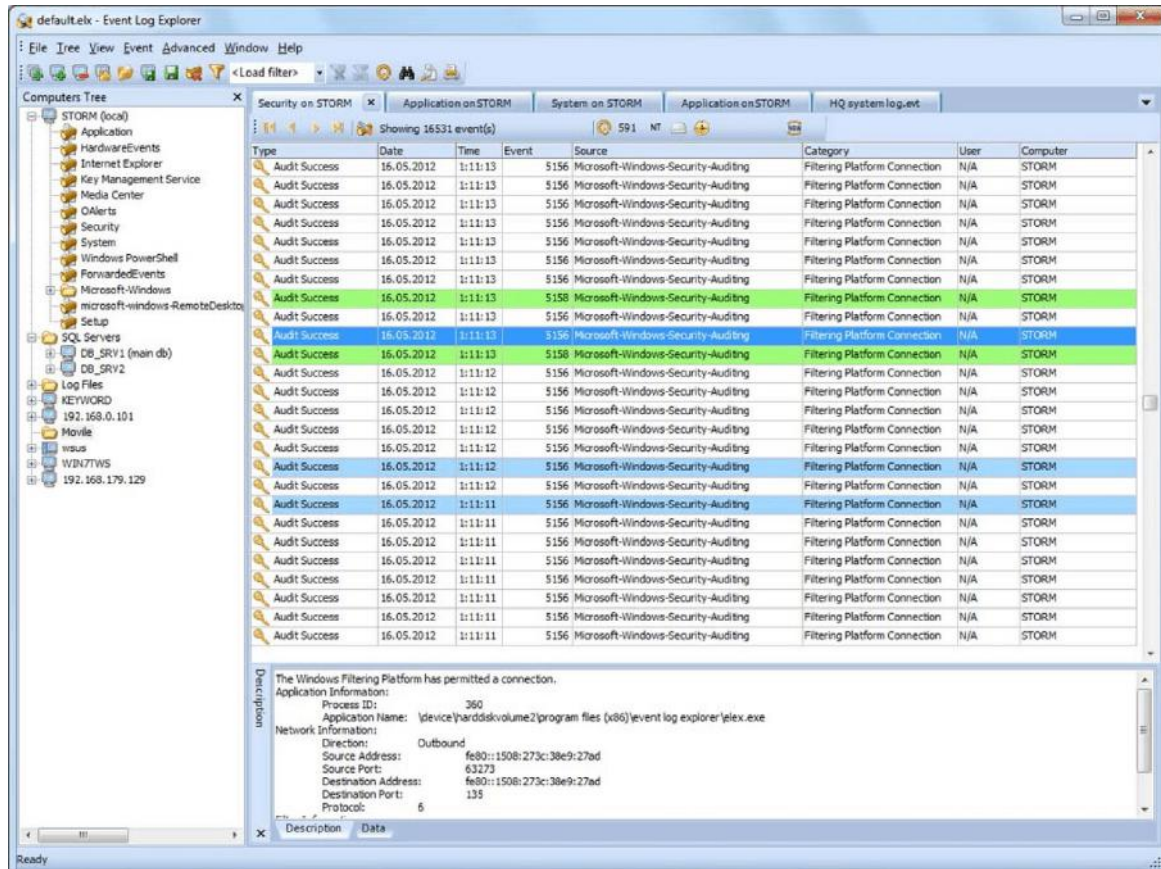
Στη συνέχεια μπορεί να αποτελείται και από πιο δευτερεύουσες λεπτομέρειες όπως:

→ Υπεύθυνος (resource) : το άτομο ή ο οργανισμός που έχει αναλάβει μία διεργασία και πρέπει να την φέρει εις πέρας.Κάθε διαδικασία/υπόθεση έχει ένα υπεύθυνο για αυτήν.

→ Κατάσταση ολοκλήρωσης : αν μία διαδικασία έχει ολοκληρωθεί επιτυχώς ή ανεπιτυχώς,αν έγινε διακοπή της λειτουργίας της ή αν ακυρώθηκε τελείως.

→ Δαπάνες : το χρηματικό ποσό που ξοδεύτηκε για να εκτελεστεί μία υπόθεση και οι διεργασίες της,συνήθως ορίζεται από την αρχή αλλά μπορεί να μεταβληθεί στη πορεία [11]

Αξίζει να σημειωθεί ότι όλα τα σύγχρονα συστήματα που χρησιμοποιούμε διαθέτουν τέτοιου είδους αρχεία.Για παράδειγμα τα Microsoft Windows,με πάνω από ένα δισεκατομμύρια χρήστες κάθε μήνα,έχουν τέτοια αρχεία καταγραφής,τα οποία είναι διαθέσιμα για τους χρήστες με σκοπό να τα μελετήσουν και σε περίπτωση κάποιου προβλήματος να διαπιστώσουν τι συνέβη.Παρόμοια εργαλεία συναντώνται και σε άλλες πλατφόρμες,αποσκοπώντας στην διερεύνηση συμβάντων και ανακάλυψη της ρίζας του προβλήματος.Στις περισσότερες περιπτώσεις αυτά λειτουργούν αυτοματοποιημένα,με το πάτημα ενός μόλις κουμπιού,παρουσιάζοντας στον χρήστη όλα τα γεγονότα του συστήματος του.



Εικόνα 11: Εικόνα από Event Log Explorer με συμβάντα στο λειτουργικό σύστημα Microsoft Windows

2.6.2 ΠΡΟΤΥΠΙΑ ΑΡΧΕΙΩΝ

Όσον αφορά το παράδειγμα της διπλωματικής τα log files διαθέτουν το XES Format (ένα XML πρότυπο για event log) που έχει ως σκοπό να προσφέρει ένα γενικό πρότυπο πληροφοριών για την παρουσίαση και επεξεργασία δεδομένων σε εργαλεία και εφαρμογές.

Το XES (extensible event stream) παρουσιάστηκε το 2010 και είναι ο πετυχημένος διάδοχος του MXML προτύπου, όπου εμφανίστηκε το 2004. Η MXML είναι μια γλώσσα δημιουργίας διεπαφής χρήστη (user interface) που βασίζεται σε XML που εισήχθη για πρώτη φορά από τη Macromedia τον Μάρτιο του 2004. Οι προγραμματιστές εφαρμογών χρησιμοποιούν τη MXML σε συνδυασμό με το ActionScript για την ανάπτυξη εμπλουτισμένων εφαρμογών web, με προϊόντα όπως το Apache Flex. Χρησιμοποιώντας το MXML είναι δυνατή η αποθήκευση αρχείων καταγραφής συμβάντων χρησιμοποιώντας μια σύνταξη βασισμένη σε XML. Το MXML έχει ένα τυπικό πρότυπο για την αποθήκευση χρονικών ετικετών, πόρων και κατάστασης της διαδικασίας. Επιπλέον, μπορεί κανείς να προσθέσει αυθαίρετα στοιχεία δεδομένων σε υποθέσεις και διαδικασίες.

Το πρότυπο XES ορίζει μια γραμματική για μία γλώσσα που βασίζεται σε ετικέτες, βασικός σκοπός της οποίας είναι να παρέχει στους σχεδιαστές πληροφοριακών συστημάτων τεχνικές και εργαλεία

για την καταγραφή συμπεριφορών συστημάτων μέσω αρχείων καταγραφής συμβάντων. Ένα σχήμα XML που περιγράφει τη δομή ενός αρχείου καταγραφής/ροής συμβάντων XES. Επιπλέον, μια βασική συλλογή από τα λεγόμενα πρωτότυπα επέκτασης XES που παρέχουν σημασιολογία σε ορισμένα χαρακτηριστικά όπως καταγράφονται στο αρχείο καταγραφής/ροής συμβάντων περιλαμβάνεται σε αυτό το πρότυπο [12]

Standard – XES (eXtensible Event Stream)

The image shows a snippet of XES XML code with several annotations in yellow boxes:

- extensions loaded:** Points to the `<extension>` declarations at the top of the code.
- every trace has a concept:name:** Points to the `<string key="concept:name" value="UNKNOWN"/>` element within the `<global scope="trace">` block.
- every event has a concept:name, lifecycle:transition and time:timestamp:** Points to the `<string key="lifecycle:transition" value="UNKNOWN"/>`, `<string key="concept:name" value="UNKNOWN"/>`, and `<date key="time:timestamp" value="1970-01-01T01:00:00+01:00"/>` elements within the `<global scope="event">` block.
- classifier name and attributes keys (concept:name, lifecycle:transition):** Points to the `<string key="concept:name" value="UNKNOWN"/>` and `<string key="lifecycle:transition" value="UNKNOWN"/>` elements within the `<global scope="event">` block.

Εικόνα 12: Δομή ενός XES αρχείου

Στις παρακάτω εικόνες παρουσιάζεται η δομή των XES αρχείων και το κάθε στοιχείο που μπορούμε να αντλήσουμε, παρατηρώντας το σκελετό του αρχείου. Στη συνέχεια παρουσιάζεται και η δομή των XES αρχείων των δήμων μέσα από το Prom Tools.

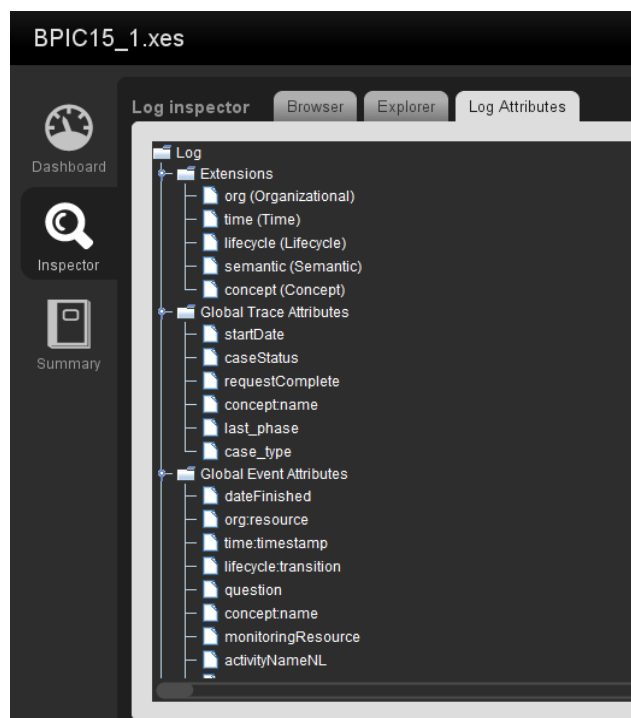
Standard – XES (eXtensible Event Stream)

The image shows a snippet of XES XML code representing a trace, with annotations in yellow boxes:

- attributes:** Points to the `<string key="concept:name" value="trace 0"/>` element within the `<trace>` block.
- event:** Points to the `<event>` blocks, each containing `<date key="time:timestamp" value="..."/>`, `<string key="concept:name" value="..."/>`, and `<string key="lifecycle:transition" value="complete"/>` elements.
- trace:** Points to the entire `<trace>` block.

Εικόνα 13: Δομή ενός ίχνους με διαδικασίες και attributes

Παραπάνω βλέπουμε αυτό που είχα προαναφέρει,ότι σε κάθε αρχείο έχουμε πολλά ίχνη ή άτομα(traces) όπου με τη σειρά του,το καθένα από αυτά περιλαμβάνει πολλές διαδικασίες(events).Κάθε event διαθέτει αρκετά στοιχεία (attributes) για να μπορεί να διαχωριστεί και να γίνει συγκρίσιμο με άλλα event.Να σημειωθεί επίσης ότι στην αρχή φορτώνονται τα extension που απαιτούνται για την ορθή λειτουργία του αρχείου.Έπειτα αναφέρονται τα στοιχεία του κάθε trace και μετά τα στοιχεία που έχει κάθε event.Τέλος αναφέρονται και κάποια classifiers ,λειτουργώντας ως κριτήρια για την σύγκριση 2 ή περισσότερων event ή traces.



Εικόνα 14: Δομή των log file του παραδείγματος μας

Στα αρχεία που μελετήσαμε στην διπλωματική έχουμε το παραπάνω πρότυπο που φαίνεται μέσα από το Inspector του Prom,στην αρχή φορτώνονται τα extension,μετά βλέπουμε τα στοιχεία των trace,τα στοιχεία των event και τέλος τα classifiers που συγκρίνουν τα δεδομένα μεταξύ τους.

Με βάση τα παραπάνω και στη συνέχεια με την επόμενη φωτογραφία μπορούμε να δούμε την λίστα με όλα τα attributes που έχουμε στα αρχεία καταγραφής συμβάντων,τόσο για τις υποθέσεις (cases) όσο για τις διαδικασίες που περιλαμβάνουν (events).Έτσι κατανοούμε καλύτερα τις σχέσεις μεταξύ των ιχνών και των διαδικασιών και πως επιδρούν μεταξύ τους.

variable	description
Case ID	a unique case number,
Activity	a short description in English
Resource	a number representing an employee; 72 different persons
Complete Timestamp	a date, and sometimes a time of completion of this activity
(case) IDofConceptCase	often empty
(case) Includes subCases	yes/no, often empty
(case) Responsible actor	a number representing an employee; 64 different persons
(case) SUMleges	the amount of money paid for the permitapplication, often empty
(case) caseProcedure	normal or extended, often empty
(case) caseStatus	Open (O) or Closed (G)
(case) case_type	a number that is the same for all activities in the log
(case) landRegisterID	a number, often empty
(case) last_phase	the last activity in the log for the case
(case) parts	a short description of the type of permit (building, demolishing, advertising,etc)
(case) requestComplete	true or false
(case) termName	not explained, often empty
action_code	a code for the type of activity, indicating position of the activity in the order of the process, eg 01 HOOFD_250_1
activityNameNL	a short description of the activity in Dutch
concept:name	not explained
dateFinished	a date, and sometimes a time of completion of the last activity for this case
dateStop	not explained, usually empty
dueDate	not explained, usually empty
lifecycle:transition	technical formality, Complete, to relate timestamp to the end of the activity
monitoringResource	a number representing an employee; 79 different persons

Εικόνα 15: Λίστα με όλα τα attributes και την ερμηνεία τους του παραδείγματος μας

Βλέπουμε ότι κάθε case και κάθε event έχουν κάποια βασικά χαρακτηριστικά όπως ID(αναγνωριστικό),name,resource(υπάλληλος),ημερομηνία έναρξης και λήξης αυτού,με σκοπό να μπορούμε να τα παρατηρήσουμε και να τα συγκρίνουμε μεταξύ τους για να βγάλουμε βασικά συμπεράσματα για την συμπεριφορά τους.Ακολουθούν και άλλα χαρακτηριστικά όπως αν είναι ολοκληρωμένη ή όχι,ποιός οργανισμός παρακολουθεί την εξέλιξη της,το χρηματικό ποσό που απαιτείται και άλλα.Γενικά κάθε διαδικασία αποτελείται από μία υπόθεση,μία ονομασία,μία χρονοσφραγίδα και όποια άλλα στοιχεία υπάρχουν.

2.7 ΠΡΟΚΛΗΣΕΙΣ ΣΤΗΝ ΤΟΠΙΚΗ ΑΥΤΟΔΙΟΙΚΗΣΗ

Βέβαια όπως είναι γνωστό ο δημόσιος τομέας και η τοπική αυτοδιοίκηση δεν εντάσσονται κατά γράμμα στις επιχειρήσεις που ξέρουμε, με στόχο το κέρδος μέσα από αγαθά και υπηρεσίες. Το βασικό τους μέλημα, μέσα από τις κατευθυντήριες γραμμές της πολιτείας, είναι να εκτελούν εργασίες και να προσφέρουν υπηρεσίες, για να βελτιώσουν το βιοτικό επίπεδο των πολιτών και να εξυπηρετούν τις βασικές ανάγκες τους. Για αυτόν τον ιδιαίτερο λόγο όταν συζητάμε για την εξόρυξη διαδικασιών στον δημόσιο τομέα υπάρχουν κάποιες προκλήσεις που θα πρέπει να λάβουμε υπόψη.

Ένα σοβαρό ζήτημα στους δήμους είναι οι χρόνοι εκτέλεσης εργασιών αλλά και οι οργανισμοί που τις εκτελούν.

Θα πρέπει η εξόρυξη διαδικασιών μέσα από τα εργαλεία της να είναι σε θέση να μειώσει τους χρόνους και τα κόστη που απαιτούνται για την διεκπεραίωση εργασιών για τους δήμους. Μέσα από την ανάλυση και την μελέτη των εργασιών μπορούμε να διαπιστώσουμε ποιες εργασίες είναι αδικαιολόγητα χρονοβόρες και επιζήμιες και να τις τροποποιήσουμε. Ακόμα με τα παραπάνω εργαλεία μπορεί να φανεί ποιοι οργανισμοί-υπάλληλοι δεν αποδίδουν ικανοποιητικά και με την κατάλληλη διαχείριση αυτών, να αυξηθεί η παραγωγικότητα.

Μία ακόμα πρόκληση είναι η εύρεση, συγχώνευση και μελέτη των event log. Κατά την εξαγωγή δεδομένων κατάλληλα για εξόρυξη διεργασιών, πρέπει να αντιμετωπιστούν διάφορες προκλήσεις: τα δεδομένα μπορεί να έχουν χωριστεί σε διάφορες πηγές, τα δεδομένα μπορεί να έχουν κενά, ένα αρχείο log μπορεί να περιέχει ανακριβείς τιμές, τα αρχεία μπορεί να περιέχουν πληροφορίες που προέρχονται από διαφορετικά επίπεδα.

Ακόμα μία πρόκληση είναι η δημιουργία κατάλληλων αντιπροσωπευτικών συγκριτικών αξιολογήσεων (benchmark), αποτελούμενα από ορθολογικά δεδομένα και συμπεράσματα, για να βοηθήσουν στην μελέτη και βελτίωση των εργαλείων και αλγορίθμων που διαθέτουμε.

Ακόμα μία σημαντική πρόκληση που θα πρέπει να διευθετηθεί μέσα από την εξόρυξη διαδικασιών είναι η εξισορρόπηση μεταξύ ποιοτικών κριτηρίων όπως η φυσική κατάσταση (fitness), η απλότητα (simplicity), η ακρίβεια (precision) και η γενίκευση (generalization), αυτά αποτελούν τέσσερις ανταγωνιστικές ποιοτικές διαστάσεις. Η πρόκληση είναι να βρεθούν μοντέλα διαδικασιών που βαθμολογούν καλά και στις τέσσερις διαστάσεις και όχι μόνο σε μία ή δύο από αυτές. Αυτό δείχνει και κατά πόσο είναι ολοκληρωμένη η εικόνα που παρουσιάζεται.

Όπως αντιλαμβανόμαστε όλοι, οι δήμοι δεν είναι αποκωμένοι ο ένας από τον άλλον, κάτι τέτοιο βέβαια δεν θα ήταν συνετό όσον αφορά την αλληλοβοήθεια και την ταχύτερη διεκπεραίωση υποχρεώσεων από αυτούς. Στην πραγματικότητα όπως και στην δικιά μας περίπτωση οι δήμοι μοιράζονται κοινές εργασίες και υποθέσεις, ενώ όπως φαίνεται και από τις μετρήσεις παρακάτω, πολλοί υπάλληλοι εργάζονται σε παραπάνω από ένα δήμους. Αυτό καθιστά μία πρόκληση για την εξόρυξη διεργασιών, καθώς θα πρέπει να λάβει σοβαρά υπόψιν αυτά τα ζητήματα και να τα διαχειριστεί κατάλληλα.

Συμπληρώνοντας ακόμα το ουσιαστικότερο ζήτημα της εξόρυξης διεργασιών είναι μετά από όλες τις αναλύσεις, τις μετρήσεις και τα συμπεράσματα να είναι σε θέση να βελτιώσει την διαδικασία, ακόμα και για τους μη εξοικιωμένους χρήστες. Θα πρέπει μέσα από αξιόπιστα εργαλεία και αλγόριθμους φιλικούς προς το χρήστη, να βελτιώσει την χρησιμότητα και κατανόηση των πλατφορμών εξόρυξης, με σκοπό να τις διαχειρίζονται με ευκολία και εμπιστοσύνη όλοι οι χρήστες.
[12]

ΚΕΦΑΛΑΙΟ 3

3.1 ΜΕΘΟΔΟΛΟΓΙΑ ΕΞΟΡΥΞΗΣ ΔΙΑΔΙΚΑΣΙΩΝ

Για να γίνουν πλήρως κατανοητά τα βήματα που ακολουθήθηκαν στο παράδειγμα της διπλωματικής, θα παρουσιαστούν στη συνέχεια αναλυτικά τα βήματα τα οποία έγιναν. Στο πρώτο σκέλος θα θέλαμε να παρουσιάσουμε σε ένα πιο γενικό βήμα τις οδηγίες που θα έπρεπε να ακολουθηθούν, σε περιπτώσεις που επιδιώκουμε την εξόρυξη διαδικασιών σε μία οποιαδήποτε περίπτωση. Τα βήματα παρουσιάζουν ένα γενικευμένο μονοπάτι της εξόρυξης διαδικασιών που θα είναι δυνατό να εφαρμοστεί σε πολλές διαφορετικές περιπτώσεις. Είτε έχουμε τις διαδικασίες ενός δήμου ή μίας εταιρίας, είτε έχουμε τις ενέργειες ενός ατόμου στην καθημερινότητα του, η εξόρυξη διαδικασιών με τα εργαλεία της μπορεί να μας βοηθήσει να μελετήσουμε τις συμπεριφορές, να βρούμε τα επαναλαμβανόμενα μοτίβα και να τα μελετήσουμε. Αυτές οι ενέργειες αφορούν οποιαδήποτε πλατφόρμα και εργαλεία χρησιμοποιούμε και μπορούν κάλλιστα να εφαρμοστούν σε οποιαδήποτε περιβάλλον βρισκόμαστε για εξόρυξη διαδικασιών. Στη συνέχεια θα παρουσιαστεί ο κύριος οδηγός της εξόρυξης που μπορεί να ακολουθηθεί σε όποια κατάσταση και αν βρισκόμαστε και θέλουμε να αναλύσουμε τις διαδικασίες που συμβαίνουν γύρω μας.

Import file

Το πρώτο βήμα σχεδόν πάντα αφορά την εισαγωγή των αρχείων καταγραφής συμβάντων που διαθέτουμε και θέλουμε να αναλύσουμε. Τα αρχεία περιλαμβάνουν διαδικασίες που έλαβαν χώρα σε μία εταιρία, έναν οργανισμό ή στην ζωή ενός ατόμου από το πρωί μέχρι το βράδυ. Τα δεδομένα συλλέγονται από διάφορες πηγές της καθημερινότητας μας και διαθέτουν κάποια βασικά χαρακτηριστικά, όπως ένα αναγνωριστικό, ένα όνομα και μία χρονική στιγμή που εκτελέστηκε. Ακόμα μπορούν να διαθέτουν και άλλες ιδιότητες όπως αν έχει ολοκληρωθεί η διαδικασία και άλλα. Εφόσον έχουμε συγκεντρωμένα τα αρχεία εισόδου πρέπει να τα εισάγουμε στην επιθυμητή πλατφόρμα που χρησιμοποιούμε. Για να γίνει αυτό όμως θα πρέπει να είναι σε κατάλληλο πρότυπο που να υποστηρίζει το λογισμικό που έχουμε. Συνήθως στα αρχεία καταγραφής συμβάντων έχουμε τα πρότυπα XML, XES, MXML ή CSV που υποστηρίζουν την καταγραφή γεγονότων και αντίστοιχα οι πλατφόρμες διαχειρίζονται τέτοιου είδους αρχεία. Αν το αρχείο μας έχει κάποια άλλο πρότυπο μπορούμε να χρησιμοποιήσουμε, αν διαθέτει βέβαια η πλατφόρμα μας, κάποιο πρόσθετο για την μετατροπή του αρχείου στο επιθυμητό πρότυπο. Αφού ολοκληρωθεί το πρώτο βήμα έχουμε το αρχείο εισόδου φορτωμένο στην πλατφόρμα που χρησιμοποιούμε και έτοιμο να το αναλύσουμε και να το διαχειριστούμε.

Event log visualization

Εφόσον έχουμε το αρχείο έτοιμο στο λογισμικό μας, το επόμενο βήμα θα είναι να παρατηρήσουμε και να αναλύσουμε, τα στοιχεία που μας προσφέρει η εικονικοποίηση των δεδομένων. Με αυτό το βήμα παρουσιάζονται όλα τα δεδομένα του αρχείου καταγραφής συμβάντων όπως τα ίχνη που υπάρχουν, οι διαδικασίες που περιλαμβάνουν και άλλα στοιχεία. Κατά την εικονικοποίηση δεδομένων μπορούμε να έχουμε μία πρώτη ματιά στα αρχεία και να μελετήσουμε τις πρώτες

πληροφορίες που αντλούμε. Με αυτόν τον τρόπο έχουμε την ικανότητα να μελετήσουμε την συμπεριφορά και τα μοτίβα που συναντάμε και να τα κατανοήσουμε πλήρως. Στη συνέχεια μπορούμε με τα δεδομένα που βρήκαμε να επεξεργαστούμε τα αρχεία κατάλληλα. Κάθε πλατφόρμα εξόρυξης διαδικασιών προσφέρει πληθώρα τεχνικών εικονικοποίησης ανάλογα με τα δεδομένα που θέλουμε να αντλήσουμε. Υπάρχουν τεχνικές που να μας δείχνουν μία στατιστική εικόνα του αρχείου με τον αριθμό των ιχνών και διαδικασιών, τον μέσο όρο αυτών και τις ημερομηνίες αρχής και τέλους του αρχείου. Μπορούμε αν θέλουμε να βρούμε κάποια συχνά επαναλαμβανόμενη εργασία ή σπάνια εργασία για να βρούμε μία ανομοιογένεια και να μελετήσουμε την συμπεριφορά της. Ακόμα γίνεται να βρούμε την συσχέτιση των γεγονότων συναρτήσει του χρόνου για να δούμε πως συμπεριφέρονται σε βάθος χρόνου σε διαγράμματα για πλήρη κατανόηση. Σε κάθε περίπτωση ανεξάρτητα από την πληροφορία που ψάχνουμε, οι τεχνικές εικονικοποίησης προσφέρουν την δυνατότητα της μελέτης και ανάλυσης των πληροφοριών, όπου αργότερα θα μας βοηθήσουν να εξάγουμε συμπεράσματα και να τα επεξεργαστούμε. Τέλος όλες αυτές οι τεχνικές εφαρμόζονται χωρίς καμία επεξεργασία ή φιλτράρισμα του αρχείου που να ζητείται από το χρήστη, ωστόσο με την χρήση αυτών των εργαλείων διευκολύνεται το μετέπειτα φιλτράρισμα που πρέπει να γίνει, καθώς χρησιμοποιεί στοιχεία που έχουν εξαχθεί από την εικονικοποίηση.

Discover process model

Ένα ακόμα βήμα που πρέπει να ακολουθηθεί είναι η δημιουργία του μοντέλου διαδικασιών, κατά το οποίο παρουσιάζονται όλες οι διαδικασίες που εκτελέστηκαν στο αρχείο καταγραφής συμβάντων. Το μοντέλο διαδικασιών περιγράφει τη ροή της εργασίας ή των δραστηριοτήτων, συνήθως σε γραφική μορφή, που συμβάλλουν στην επίτευξη ενός συγκεκριμένου στόχου. Τα μοντέλα διαδικασίας χρησιμοποιούνται συνήθως για την αναπαράσταση και ανάλυση μιας σειράς δραστηριοτήτων που συμβαίνουν επανειλημμένα και σε τακτική βάση. Κάθε μοντέλο ή όπως συναντάτε συχνότερα ροή εργασιών, ξεκινάει από μία διαδικασία (start event), περιλαμβάνει άλλες καθόλη τη διάρκεια και ολοκληρώνεται με μία συγκεκριμένη εργασία (end event). Για την παραγωγή αυτών των μοντέλων, ανεξάρτητα από την πλατφόρμα που δουλεύουμε, η μεθοδολογία είναι περίπου ίδια. Αφού έχουμε το αρχείο εφαρμόζουμε mining algorithms, όπως alpha miner, heuristic miner και άλλους, για να παραχθεί το μοντέλο που ψάχνουμε. Αυτοί οι αλγόριθμοι διαφέρουν μεταξύ τους και χρησιμοποιούμε όποιον χρειαζόμαστε σε κάθε πρόβλημα, ενώ υπάρχουν πάρα πολλοί τέτοιοι σε όλες τις πλατφόρμες εξόρυξης διαδικασιών. Με την εφαρμογή τέτοιων αλγόριθμων στη πλατφόρμα που χρησιμοποιούμε, θέτουμε σαν είσοδο το αρχείο καταγραφής συμβάντων και παράγουμε σαν έξοδο ένα μοντέλο με όλες τις διαδικασίες που υπάρχουν. Με αυτόν τον τρόπο παράγουμε μία αναπαράσταση, ακόμα και ζωντανή σε πραγματικό χρόνο, των διαδικασιών του αρχείου, έχοντας την δυνατότητα να τις μελετήσουμε, να εντοπίσουμε σφάλματα ή ασάφειες που δεν είναι ορατές, βλέποντας το αρχείο και να τις διορθώσουμε. Ακόμα μπορούμε να δούμε κατά πόσο συμφωνεί το αρχικό αρχείο που έχουμε με το μοντέλο διαδικασιών που παράγεται, εντοπίζοντας αποκλίσεις. Αρκετές φορές επιλέγουμε να εκτελέσουμε πρώτα το φιλτράρισμα στο αρχείο, το οποίο είναι το επόμενο βήμα που θα παρουσιαστεί, και στη συνέχεια να εφαρμόσουμε στο επεξεργασμένο αρχείο, έναν αλγόριθμο mining. Με αυτήν την τεχνική παρατηρούμε και σε γραφικό επίπεδο πως συμπεριφέρεται το αρχείο στο φίλτρο που εφαρμόζουμε.

Filter event log

Το τελευταίο βήμα που αναφέραμε και παραπάνω και το πιο σημαντικό που εφαρμόζουμε στο αρχείο καταγραφής συμβάντων είναι το φιλτράρισμα. Μόλις εξαχθούν τα αρχεία καταγραφής συμβάντων από τα δεδομένα, το επόμενο βήμα είναι να τα φιλτράρουμε. Το φιλτράρισμα είναι μια επαναληπτική διαδικασία. Σκοπός αυτής της διαδικασίας είναι να επεξεργαστούμε τα δεδομένα με βάση κάποιο κριτήριο επιλογής που έχουμε και να παραχθεί το επιθυμητό αρχείο εξόδου. Όλες οι πλατφόρμες εξόρυξης διαδικασιών διαθέτουν μία πληθώρα από plug in, τα οποία εφαρμόζονται στο αρχείο και το φιλτράρουν κατάλληλα. Μπορούμε να ασχοληθούμε με τις υποθέσεις ή με τις διαδικασίες που αποτελούνται. Έχουμε την ευκαιρία να μελετήσουμε σημεία ενδιαφέροντος ή ανομοιογένειες στα δεδομένα, όπως συχνά επαναλαμβανόμενες διαδικασίες και υποθέσεις ή αντίστοιχα σπάνιες υποθέσεις. Εφαρμόζοντας τα προηγούμενα βήματα και εργαλεία αναλύουμε και εντοπίζουμε αποκλίσεις, σημεία συμφόρησης και σφάλματα στα αρχεία καταγραφής συμβάντων. Σε αυτό το βήμα έχοντας αντιληφθεί τέτοια ζητήματα, επεξεργαζόμαστε τα δεδομένα για να μελετήσουμε επακριβώς τα προβλήματα για να δούμε πως συμπεριφέρονται μεμονωμένα, χωρίς τον υπόλοιπο όγκο πληροφοριών. Με αυτόν τον τρόπο καταλαβαίνουμε καλύτερα πως λειτουργούν και ποία μοτίβα αναπτύσσουν. Καταλήγουμε να βρίσκουμε την ρίζα του προβλήματος και τα αίτια που τη δημιούργησαν, με επόμενο βήμα να την διορθώνουμε και να βελτιώνουμε τις διαδικασίες που εκτελούνται.

3.2 ΕΞΟΡΥΞΗ ΔΙΑΔΙΚΑΣΙΩΝ ΜΕ ΤΟ PROM TOOLS

Για να γίνουν πλήρως κατανοητά τα βήματα που ακολουθήθηκαν στο παράδειγμα των 5 Ολλανδικών δήμων, θα παρουσιαστούν αναλυτικά με οδηγίες στο Prom tools οι ενέργειες που έγιναν. Στο συγκεκριμένο παράδειγμα εφαρμόστηκαν τεχνικές και αλγόριθμοι εξόρυξης διαδικασιών στους Ολλανδικούς δήμους, ωστόσο τα ίδια βήματα με ελάχιστες τροποποιήσεις μπορούν να εφαρμοστούν ικανοποιητικά και σε άλλους δήμους, όργανα τοπικής αυτοδιοίκησης ακόμα και ιδιωτικές επιχειρήσεις. Η έρευνα έγινε με την χρήση της πλατφόρμας Prom Tools ver 6.11.

Το ProM είναι ένα επεκτάσιμο πλαίσιο που υποστηρίζει μια μεγάλη ποικιλία τεχνικών εξόρυξης διεργασιών με τη μορφή πρόσθετων. Είναι ανεξάρτητο από πλατφόρμα καθώς υλοποιείται σε Java και μπορεί να το κατεβάσετε δωρεάν. Το ProM 6 διανέμεται σε μέρη, γεγονός που προσφέρει μέγιστη ευελιξία. Αρχικά, το ProM 6 διανέμεται ως πακέτο με δυνατότητα λήψης χρησιμοποιώντας την άδεια ανοιχτού κώδικα GNU Public License (GPL). Αυτό σημαίνει ότι μπορείτε να κάνετε λήψη και εγκατάσταση του ProM 6 χωρίς περιορισμούς, αλλά οποιοδήποτε λογισμικό που χρησιμοποιεί τον πυρήνα είναι απαραίτητο να γίνει λήψη και εγκατάσταση αυτού. [13]

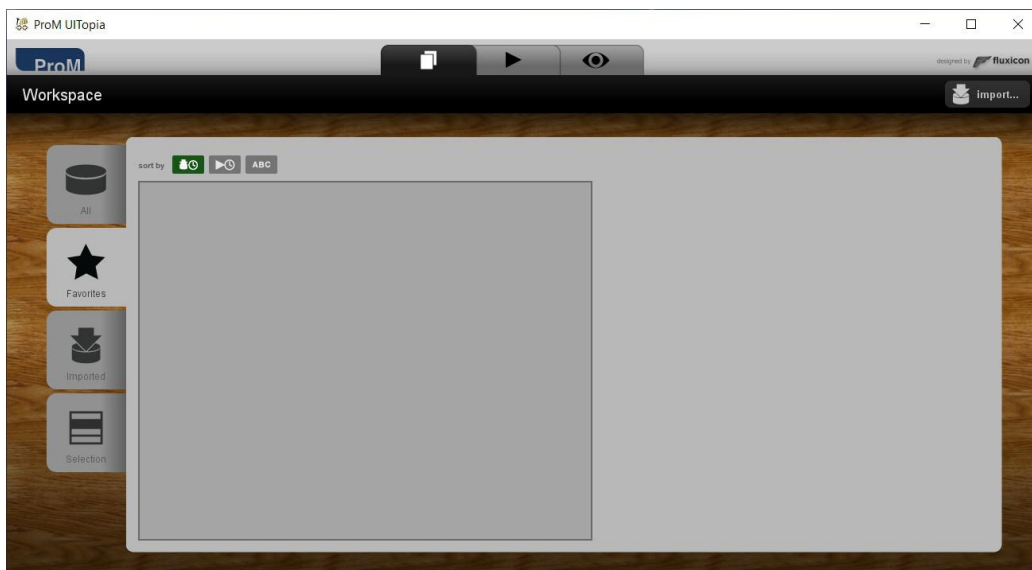
3.2.1 ΕΙΣΑΓΩΓΗ ΑΡΧΕΙΩΝ ΚΑΤΑΓΡΑΦΗΣ ΣΥΜΒΑΝΤΩΝ

Πρώτα από όλα μέσα στο εργαλείο Prom, εφόσον έχουμε την απαιτούμενη έκδοση java (jdk, jre), αφού το ανοίξουμε, βρισκόμαστε στην πρώτη καρτέλα, όπου μπορούμε στο χώρο εργασίας να

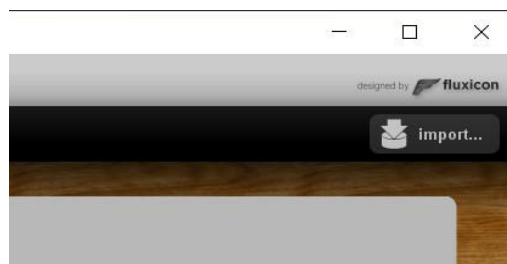
δούμε όλα τα αρχεία που εισαγάγαμε, τα αποθηκευμένα και τα αγαπημένα μας αρχεία. Έπειτα έχουμε την δυνατότητα με την επιλογή import να εισάγουμε το επιθυμητό αρχείο σε XML μορφή.

Ένα αρχείο μπορεί επίσης να είναι σε μορφή csv, mxml ή σε μορφή XES. Αν εισάγουμε ένα αρχείο σε μορφή csv, πρέπει να το μετατρέψουμε σε μορφή XES χρησιμοποιώντας Προσθήκη "Μετατροπή CSV σε XES". Σε αυτή την περίπτωση θα αντιστοιχίσουμε τις στήλες του csv στο νέο πρότυπο που προκύπτει.

Prom Workspace

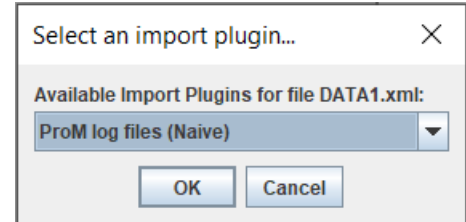
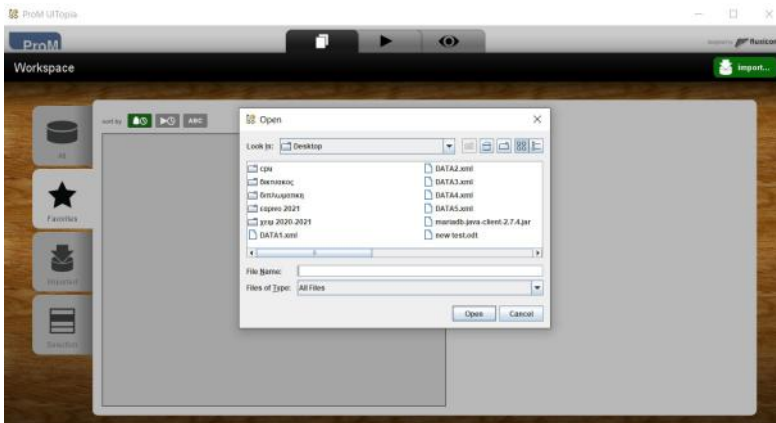


Εικόνα 16: Το περιβάλλον του Prom Tools

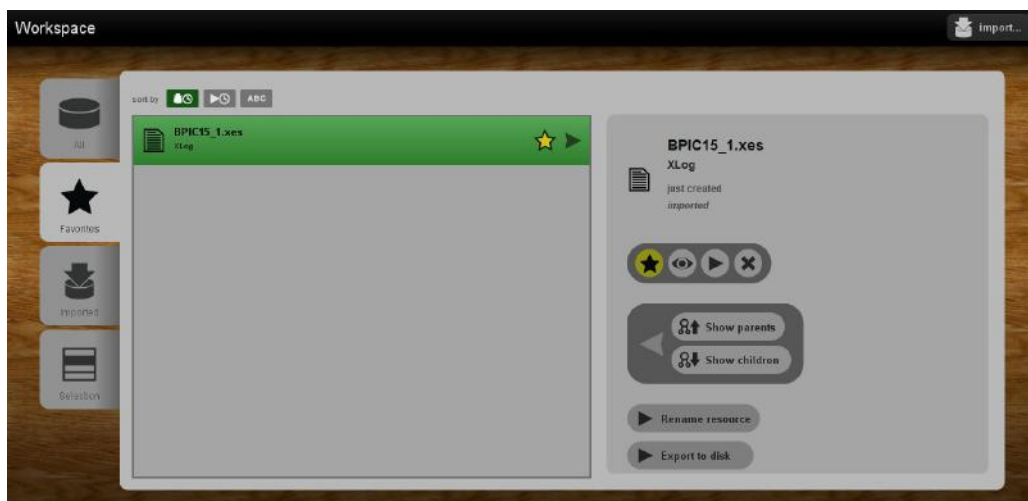


Εικόνα 17: Επιλογή για εισαγωγή αρχείου στο λογισμικό

Με την επιλογή import θα εμφανιστεί η καρτέλα με όλα τα διαθέσιμα αρχεία του δίσκου μας για να εισάγουμε όποιο θέλουμε και με όποιο από τα διαθέσιμα plug in για import χρειαζόμαστε. Το Prom σαν εργαλείο διαθέτει πάνω από 1.000 plug in προς χρήση.



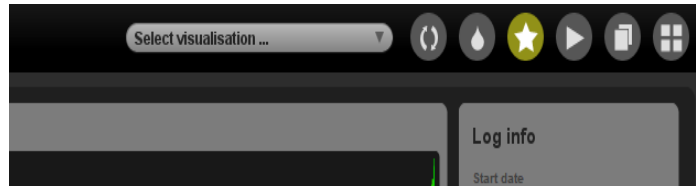
Εικόνα 18: Επιλογή του επιθυμητού αρχείου από το σύστημα μας



Εικόνα 19: Αρχική σελίδα

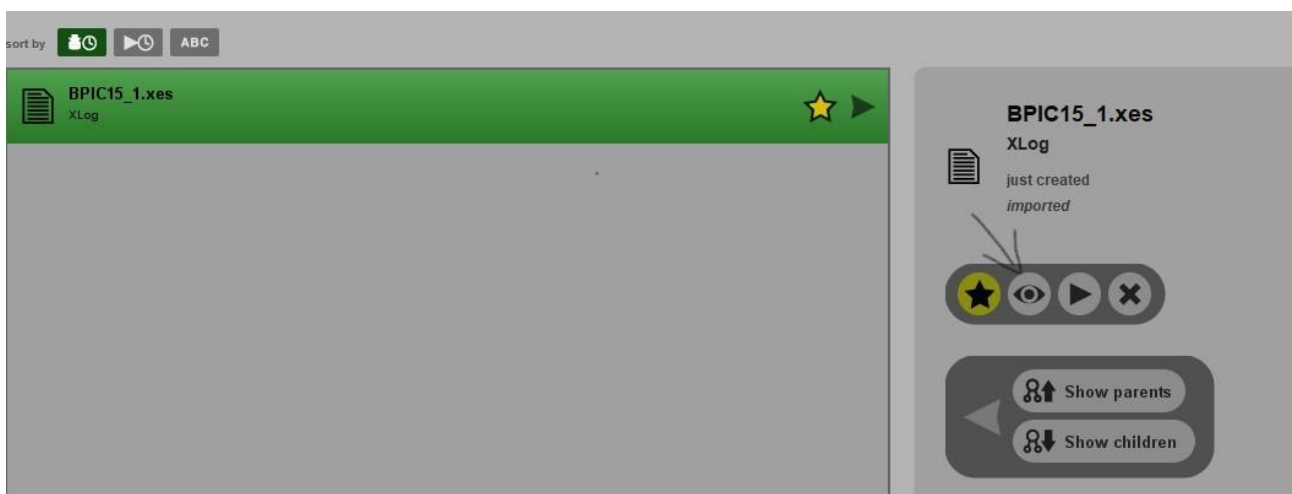
Στην καρτέλα Workspace εμφανίζεται το αρχείο που μόλις τοποθετήσαμε στην εφαρμογή και μπορούμε να το επεξεργαστούμε. Κάθε αρχείο που τοποθετούμε έχει κάποια υποχρεωτικά στοιχεία όπως αναγνωριστικό υπόθεσης (case id), όνομα διαδικασίας (event name) και χρονοσφραγίδα (timestamp). Αν διαθέτουμε και άλλες παραμέτρους τις χρησιμοποιούμε κανονικά για περαιτέρω ανάλυση και φιλτράρισμα των αρχείων.

Το επόμενο βήμα είναι να παρουσιάσουμε οπτικά τα δεδομένα και να τα μελετήσουμε από το αρχείο καταγραφής συμβάντων. Οι διαθέσιμες τεχνικές οπτικοποίησης του αρχείου βρίσκονται στην επάνω δεξιά γωνία.



Εικόνα 20: Λίστα με τις διαθέσιμες τεχνικές εικονικοποίησης των δεδομένων

3.2.2 EVENT LOG VISUALIZATION

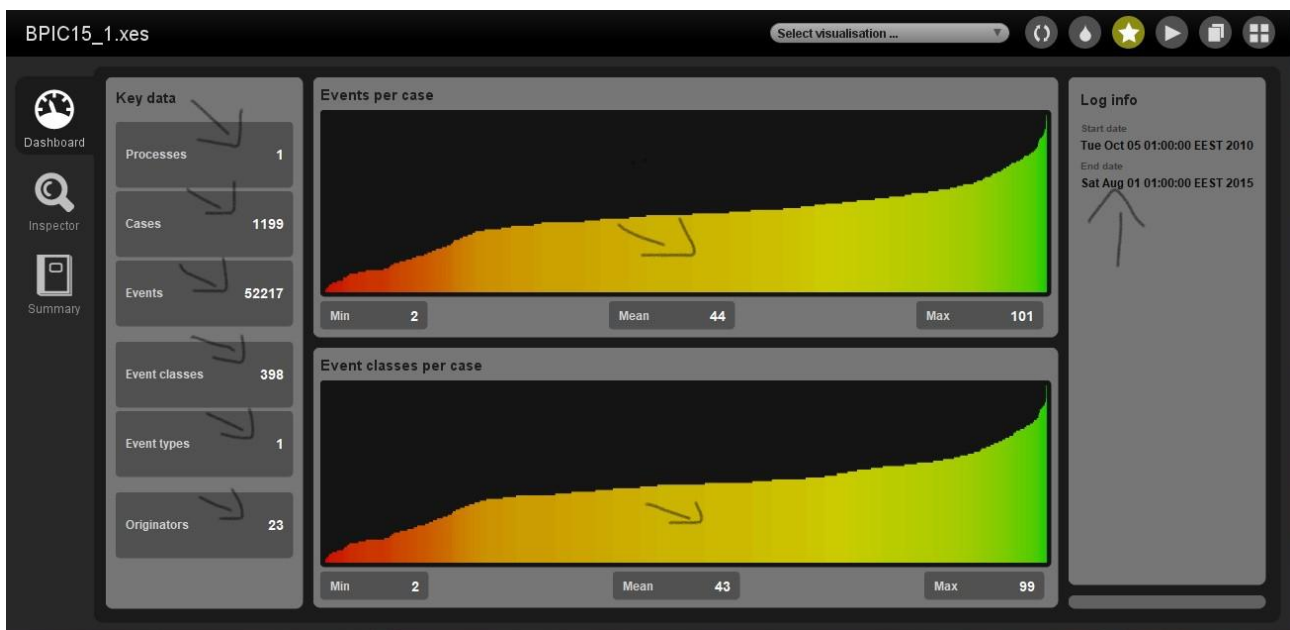


Εικόνα 21: Λίστα με τα αρχεία και των διαθέσιμων ενεργειών

Μόλις έχει γίνει αυτό το βήμα, πριν φιλτράρουμε οτιδήποτε μπορούμε να επιλέξουμε το κουμπί view resource και να παρατηρήσουμε τις πληροφορίες που μας δίνει η καρτέλα dashboard, για να βγάλουμε κάποια συμπεράσματα. Αυτό το βήμα παρουσιάζει την αναπαράσταση του αρχείου. Εν συνεχεία είμαστε σε θέση στο παράθυρο να δούμε βασικές πληροφορίες για το αρχείο όπως τον αριθμό των υποθέσεων, των δραστηριοτήτων και των οργανισμών που συμμετέχουν σε αυτά. Μέσα από την γραφική απεικόνιση βλέπουμε τις μέγιστες, ελάχιστες καθώς και το μέσο όρο, τιμών για τις διεργασίες ανά υπάλληλο.

Παρατηρούμε επίσης τον αριθμό των διαφορετικών υποθέσεων που έχουμε, καθώς και τις ημερομηνίες έναρξης και λήξης του αρχείου.

Log visualizer/dashboard



Εικόνα 22: Απεικόνιση των στοιχείων του αρχείου καταγραφής συμβάντων

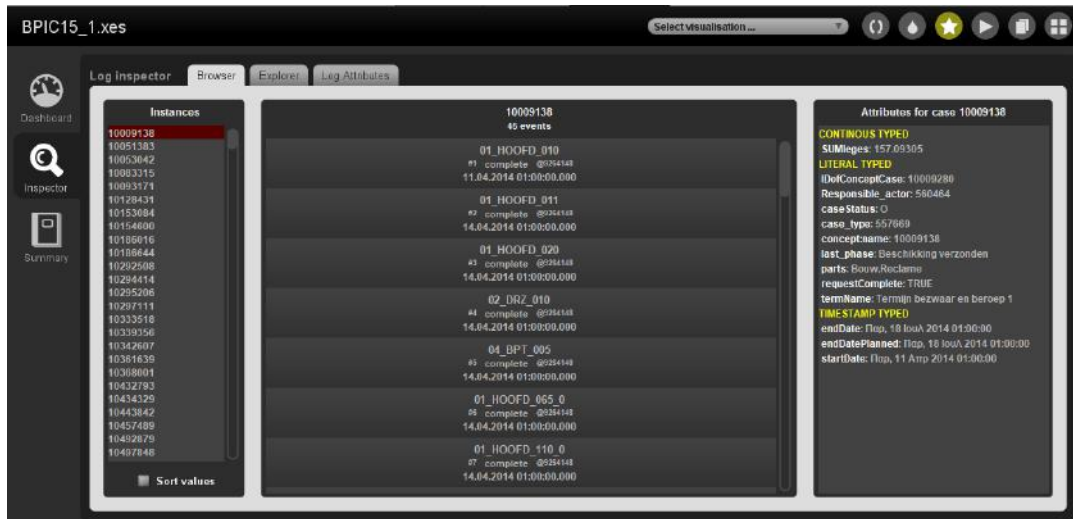
Πιο αναλυτικά από την παραπάνω εικόνα αντλούμε τις εξής πληροφορίες:

- Πόσες διεργασίες(processes) περιλαμβάνονται στο αρχείο καταγραφής συμβάντων (προφανώς μία)
- Πόσες υποθέσεις έχουμε στο αρχείο καταγραφής συμβάντων
- Πόσα γεγονότα καταγράφονται στο αρχείο καταγραφής συμβάντων (μπορεί ένα γεγονός να καταγράφεται πολλές φορές)
- Πόσες κατηγορίες/είδη γεγονότων έχουμε παρατηρήσει στο αρχείο καταγραφής συμβάντων
- Πόσοι τύποι συμβάντων (π.χ. έναρξη, ολοκλήρωση)
- Πόσοι οργανισμοί/υπάλληλοι εκτέλεσαν διεργασίες στο συγκεκριμένο αρχείο
- Πόσες διεργασίες εκτελέστηκαν ανά υπόθεση(συμπεριλαμβάνεται η μέγιστη,ελάχιστη και μέση τιμή διεργασιών ανά υπόθεση)
- Πόσες κατηγορίες διεργασιών εκτελέστηκαν ανά υπόθεση(συμπεριλαμβάνεται η μέγιστη,ελάχιστη και μέση τιμή διεργασιών ανά υπόθεση)
- Η πρώτη και η τελευταία παρατηρούμενη χρονική διεργασία, η οποία δίνει το χρονικό διάστημα το αρχείο καταγραφής συμβάντων

Inspector

Στα επόμενα παράθυρα μέσω του Inspector έχουμε 3 διαφορετικές προσεγγίσεις:browser,explorer και log attributes.Κάθε μία δίνει διαφορετικές λεπτομέρειες για τις υποθέσεις,τις διεργασίες και τους εργαζομένους.

Browser



Εικόνα 23: Καρτέλα browser στον inspector

Ο Browser μας βοηθάει να δούμε όλα τα ίχνη(και τον αριθμό αυτών) και την σειρά εκτέλεσης των γεγονότων που περιλαμβάνονται. Οι πληροφορίες για κάθε γεγονός περιλαμβάνουν το όνομα(κωδικός), την κατάσταση συναλλαγής (έναρξη ή ολοκλήρωση) και χρονοσφραγίδα(timestamp, πότε συνέβη). Ενώ συνολικά για κάθε ίχνος με τις διεργασίες του, μπορούμε να δούμε αναλυτικές πληροφορίες για αυτό.

Για παράδειγμα μπορούμε να δούμε:

CONTINUOUS TYPED:

→ SUMIpleges (πληρωθέν ποσό για την αίτηση άδειας)

LITERAL TYPED:

- IdofConceptCase (αναγνωριστικό υπόθεσης)
- Responsible actor (υπεύθυνος οργανισμός/υπάλληλος)
- caseStatus (κατάσταση υπόθεσης ανοικτή ή κλειστή)
- caseType (ένας αριθμός ίδιος για όλες τις υποθέσεις στο αρχείο)
- conceptname (αναγνωριστικό διεργασίας)
- last_phase (η τελευταία διαδικασία στο αρχείο)
- parts (περιγραφή για τον τύπο της διεργασίας π.χ. χτίσιμο, συντήρηση)
- requestComplete (ολοκλήρωση ή όχι της υπόθεσης)
- termName (πλήρη ονομασία της υπόθεσης)

TIMESTAMP TYPED:

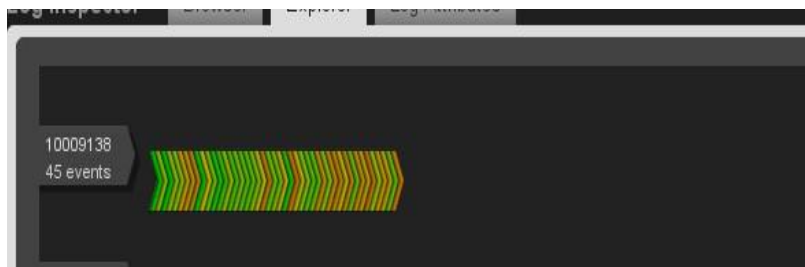
- endDate (ημερομηνία ολοκλήρωσης)
- endDatePlanned (προγραμματισμένη ημερομηνία ολοκλήρωσης)
- startDate (ημερομηνία έναρξης)

Explorer



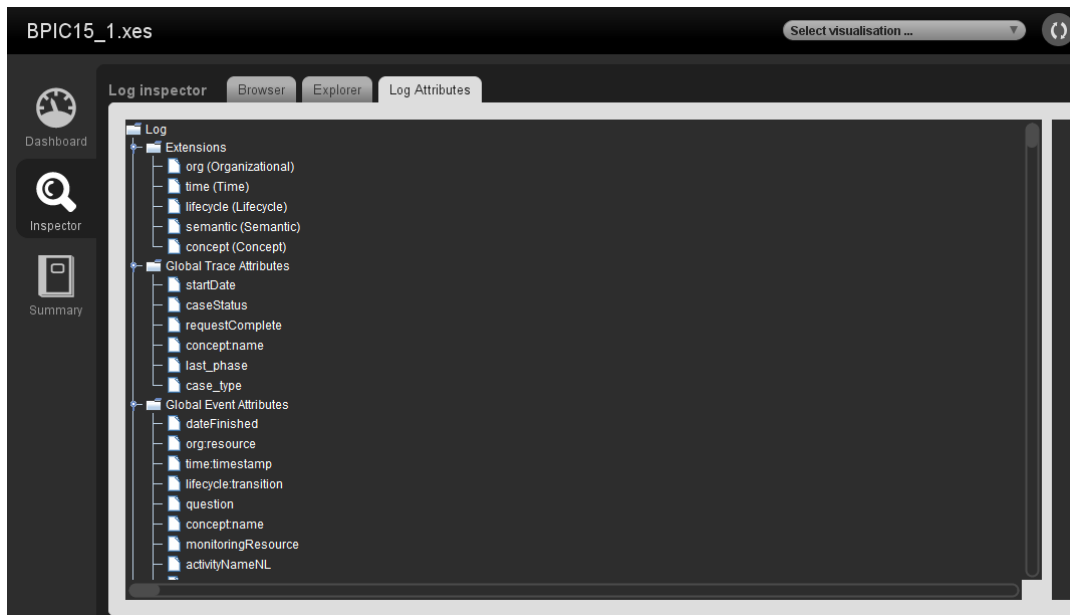
Εικόνα 24: Καρτέλα explorer στον inspector

Στην επομένη καρτέλα μπορούμε να μελετήσουμε πληροφορίες για το μήκος κάθε υπόθεσης και τα συμβάντα κάθε υπόθεσης. Το μήκος δείχνει την ποσότητα των εργασιών που εκτελούνται, ενώ αυτές χρωματίζονται με βάση τη συχνότητά εμφάνισης τους στο αρχείο. Όσο πιο συχνό το συμβάν το χρώμα είναι πιο σκούρο πράσινο, τόσο πιο σπάνιο είναι το συμβάν έχει το πιο σκούρο πορτοκαλί χρώμα. Επίσης τοποθετώντας τον κέρσορα πάνω από το γεγονός, εμφανίζονται πληροφορίες σχετικά με τον αριθμό της υπόθεσης, το όνομα της διεργασίας, την κατάσταση, τον υπάλληλο, τη συχνότητα εμφάνισης και την χρονοσφραγίδα.



Εικόνα 25: Καρτέλα explorer-στοιχεία για το ίχνος

Log attributes



Εικόνα 26: Καρτέλα Log Attributes

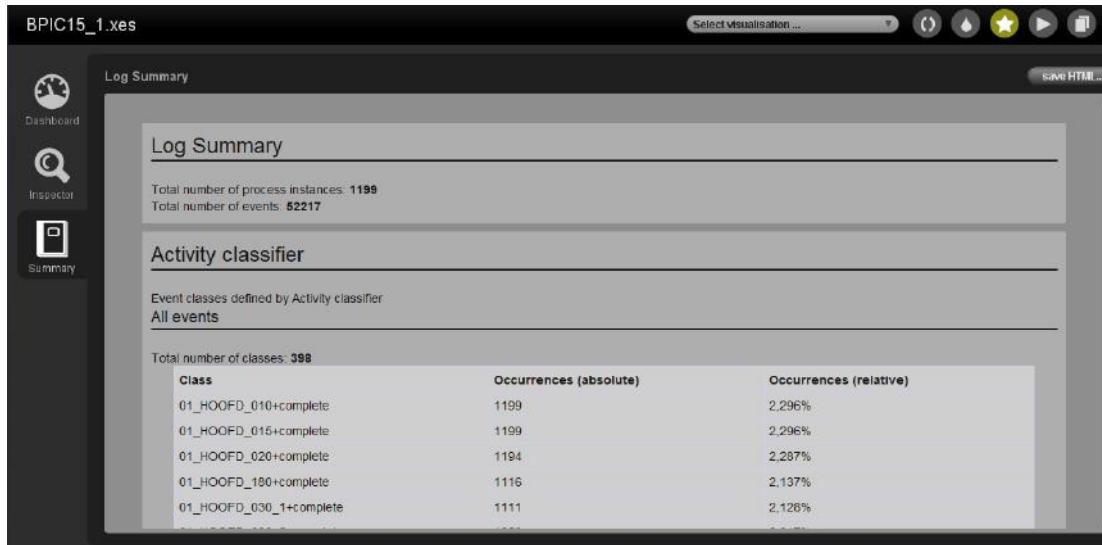
Στην καρτέλα log attributes παρέχονται πληροφορίες για την δομή του αρχείου που αφορούν τα extensions που έχουν χρησιμοποιηθεί, τα attributes τόσο των cases και των events και τα classifier που έχουν χρησιμοποιηθεί.

Παρατηρώντας την εικόνα κατανοούμε

- τα extension που έχουν φορτωθεί στο log file.
- τα attribute των υποθέσεων(trace)
- τα attribute των διαδικασιών(event)
- τους διαχωριστές (classifier) του αρχείου, με βάση τους οποίους, συγκρίνονται τα δεδομένα

Ένας διαχωριστής ή αλλιώς ταξινομητής, χρησιμοποιείται για να περιγράψει την έννοια μίας πράξης. Για παράδειγμα στα αρχεία μας έχουμε 4 ταξινομητές έναν για τον κωδικό της πράξης/ενέργειας, έναν για το όνομα της ενέργειας στα Αγγλικά και έναν άλλον για το ίδιο όνομα στην πρωτότυπη γλώσσα των αρχείων τα Ολλανδικά. Ο τελευταίος ταξινομητής αφορά τον υπάλληλο που διεκπαιρώνει την εργασία.

Summary



Start events

Total number of classes: 14

Class	Occurrences (absolute)	Occurrences (relative)
01_HOOFD_010+complete	1138	94,912%
01_HOOFD_011+complete	34	2,836%
11_AH_II_040b+complete	7	0,584%
01_HOOFD_030_2+complete	6	0,5%
01_HOOFD_065_2+complete	4	0,334%
02_DRZ_010+complete	2	0,167%
11_AH_II_010+complete	1	0,083%
05_EIND_020+complete	1	0,083%
11_AH_II_037_2+complete	1	0,083%
01_HOOFD_065_0+complete	1	0,083%
01_HOOFD_080+complete	1	0,083%
01_HOOFD_490_3+complete	1	0,083%
01_HOOFD_490_5+complete	1	0,083%
01_HOOFD_195_2+complete	1	0,083%

End events

Total number of classes: 108

Class	Occurrences (absolute)	Occurrences (relative)
01_HOOFD_530+complete	295	24,604%
01_HOOFD_510_2+complete	132	11,009%
01_HOOFD_516+complete	82	6,839%
01_HOOFD_510_2a+complete	52	4,337%
01_HOOFD_820+complete	51	4,254%
01_BB_770+complete	33	2,752%
01_HOOFD_030_2+complete	30	2,502%
01_HOOFD_519+complete	30	2,502%
01_HOOFD_530a+complete	29	2,419%

Εικόνα 27: Καρτέλα Log Summary

Στην ενότητα Summary ή αλλιώς σύνοψη μπορούμε, όπως είναι εύκολα κατανοητό, να δούμε την συνολική εικόνα του αρχείου καταγραφής συμβάντων.

Από αυτήν την καρτέλα αντλούμε πληθώρα δεδομένων με ποσοτικό χαρακτήρα για ακρίβεια.

Μπορούμε να παρατηρήσουμε:

- τον συνολικό αριθμό των υποθέσεων στο αρχείο καταγραφής συμβάντων
- τον συνολικό αριθμό των ενεργειών

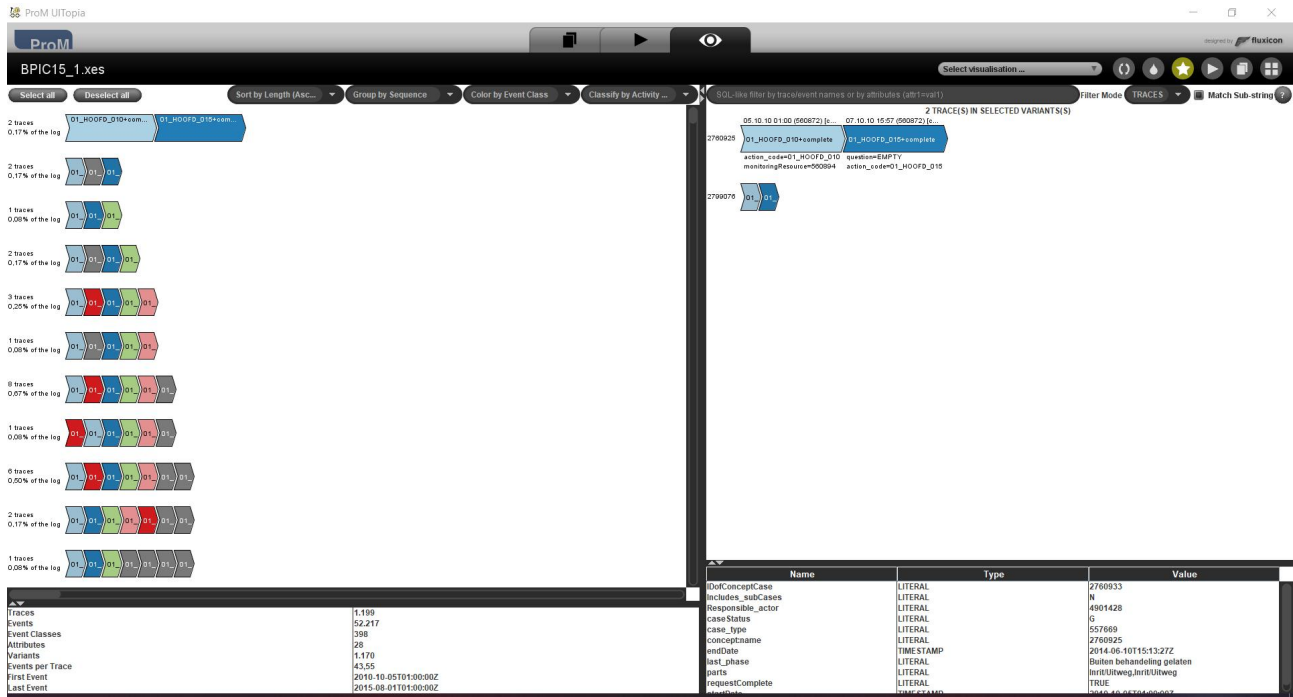
Ενώ με την χρήση των ταξινομητών για κάθε περίπτωση έχουμε:

- για κάθε γεγονός: την ονομασία του, τον αριθμό εμφανίσεων και το ποσοστό του
- για κάθε διαδικασία, τα γεγονότα με τα οποία ξεκινάει και συγκεκριμένα την ονομασία, τον αριθμό εμφανίσεων και το ποσοστό του
- για κάθε διαδικασία, τα γεγονότα με τα οποία ολοκληρώνεται και συγκεκριμένα την ονομασία, τον αριθμό εμφανίσεων και το ποσοστό του

Πρέπει να σημειωθεί ότι στις ενέργειες, στη στήλη με το όνομα τους περιλαμβάνεται και η κατάσταση της ενέργειας. Αν μετά το όνομα της ενέργειας ακολουθεί η φράση “+complete” αυτό δείχνει, ότι αυτή η διεργασία ξεκίνησε και έχει ολοκληρωθεί κανονικά. Αλλιώς στην περίπτωση που ακολουθεί η φράση “+start” υποδεικνύεται ότι ξεκίνησε αλλά δεν έχει ολοκληρωθεί ακόμα. Αυτό είναι ένα πολύ χρήσιμο εργαλείο, καθώς είμαστε σε θέση να δούμε ποιες διεργασίες στο αρχείο δεν έχουν ολοκληρωθεί ακόμα. Συνεισφέροντας στην κατανόηση του ζητήματος και εστιάζοντας στις ευπάθειες με τις οποίες έχουμε να κάνουμε κάθε φορά, αυτή η δυνατότητα παίζει μείζων ρόλο στην ανάλυση του προβλήματος. Μελετώντας την είμαστε σε θέση να εξάγουμε κάποια συμπεράσματα, τα οποία όπως θα δούμε και αργότερα με φιλτράρισμα μπορούμε να τα μελετήσουμε εις βάθος.

Explore Event Log (Trace variants/Searchable/Sortable)(Log Enhancement)

Μίας από τις πιο βασικές τεχνικές οπτικοποίησης των δεδομένων που έχουμε, είναι με την χρήση του εργαλείου explore event log. Με την χρήση αυτού είμαστε σε θέση να δούμε, ταξινομημένα με διάφορα κριτήρια όπως το μήκος ή η ακολουθία, όλα τα ίχνη με τις ενέργειες που περιλαμβάνουν. Ακόμα εμφανίζει τον αριθμό των επαναλαμβανόμενων ίχνων με ίδια event, βοηθώντας να κατανοήσουμε το τρόπο συμπεριφοράς των ενεργειών. Μπορούμε να παρατηρήσουμε τα μοτίβα που σχηματίζονται και με τι συχνότητα εμφάνισης. Τέλος για ένα αρχείο είμαστε σε θέση να κατανοήσουμε αν περιλαμβάνει ίχνη με πολλές διαδικασίες ή όχι. [14]



Εικόνα 28: Explore Event Log διάγραμμα

Μετά την ταξινόμηση βλέπουμε τα ίχνη με μικρό αριθμό διαδικασιών, όπου εκεί εμφανίζονται οι μεγαλύτερες ευπάθειες στα συστήματα και θα πρέπει να αναλυθούν περαιτέρω. Παρατηρώντας τα μπορούμε να δούμε ποία μοτίβα σχηματίζουν μεταξύ τους και πως συμπεριφέρονται, για να ανακαλύψουμε μη αναμενόμενες συμπεριφορές. Φιλτράροντας κατάλληλα είμαστε σε θέση να σχολιάσουμε την κατάσταση τους και πως δημιουργούν καθυστερήσεις (bottleneck) σε άλλες διεργασίες.



Εικόνα 29: Αναλυτικές λεπτομέρειες των υποθέσεων

Επιλέγοντας ένα μοτίβο ίχνων μας εμφανίζει πληροφορίες για τις παραμέτρους τους, ποία ίχνη αποτελούνται από τις ίδιες εργασίες καθώς και δεδομένα για την φύση των ίδιων των εργασιών.

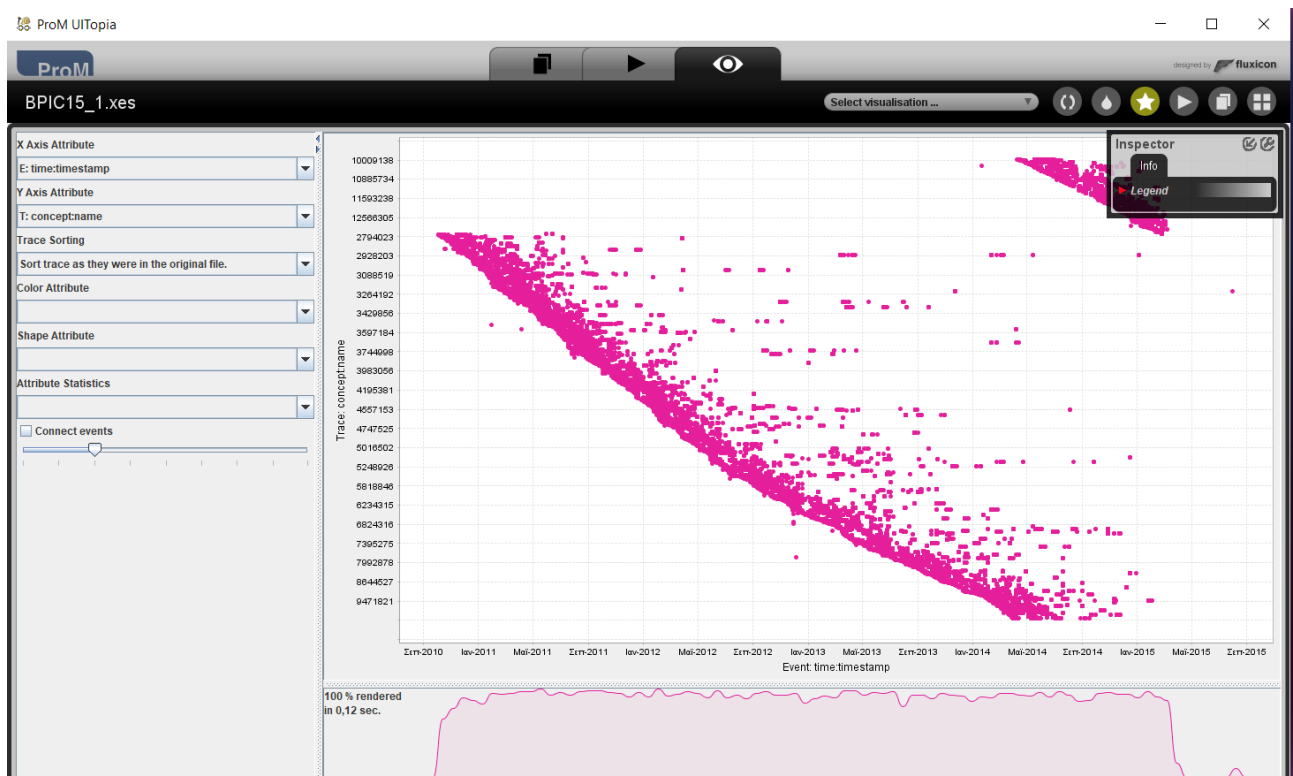
Traces	1.199
Events	52.217
Event Classes	398
Attributes	28
Variants	1.170
Events per Trace	43,55
First Event	2010-10-05T01:00:00Z
Last Event	2015-08-01T01:00:00Z

Εικόνα 30: Πληροφορίες για το αρχείο

Στο κάτω μέρος της καρτέλας περιλαμβάνονται δεδομένα για την φύση του αρχείου,όπως για παράδειγμα:

- τον αριθμό των ίχνων
- τον αριθμό των ενεργειών
- τον αριθμό των διαφορετικών κατηγοριών για τις ενέργειες που υφίστανται
- τις παραμέτρους του αρχείου τόσο για τα ίχνη όσο για τις ενέργειες
- τον αριθμό των μοτίβων που εντοπίζονται
- τον μέσο όρο ενεργειών ανά ίχνος,χρήσιμο για να δούμε αν στο αρχείο έχουμε πυκνά ίχνη
- την ημερομηνία του πρώτου event
- την ημερομηνία του τελευταίου event

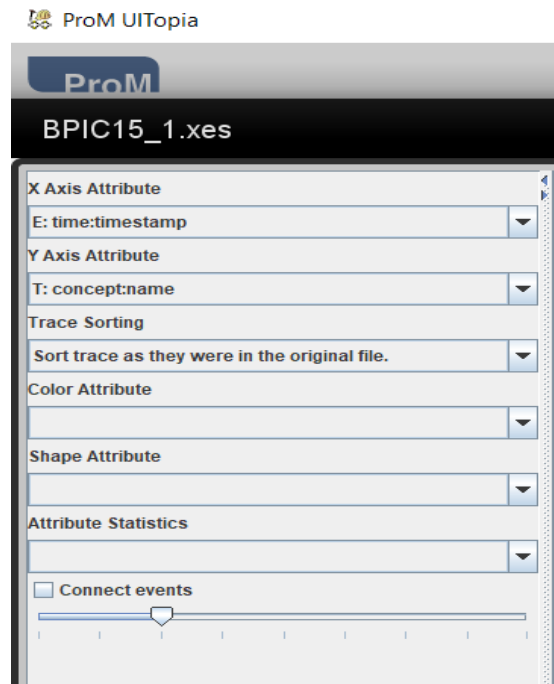
Dotted Chart



Εικόνα 31: Περιβάλλον Dotted Chart

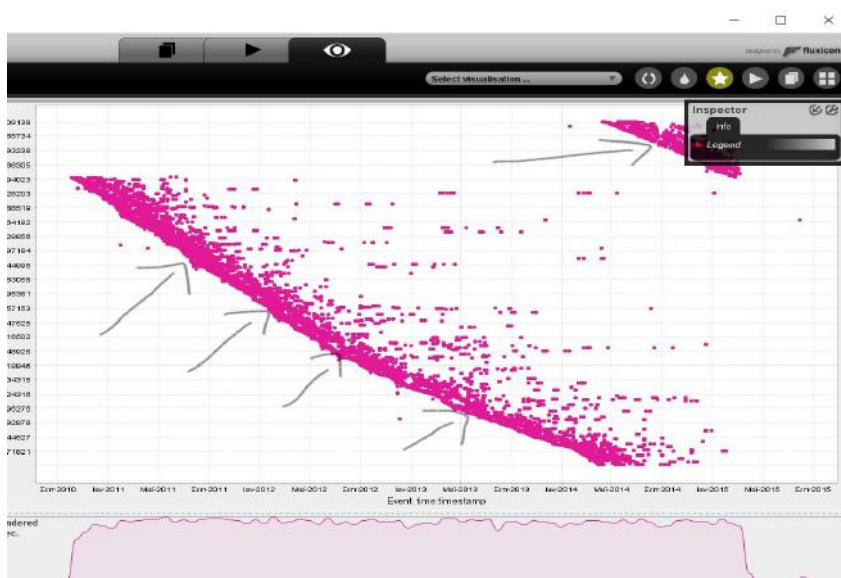
Με την αναπαράσταση αυτή είμαστε σε θέση να δείξουμε την συσχέτιση των υποθέσεων και των διαδικασιών σε συνάρτηση με το χρόνο.Είναι ιδιαίτερα χρήσιμο για την ανάλυση και εξερεύνηση

των αρχείων καταγραφής συμβάντων.Ο κάθετος και ο οριζόντιος άξονας παρουσιάζουν διάφορες μεταβλητές της επιλογής μας,όπου μπορούμε να τις παραμετροποιήσουμε ανάλογα με την έρευνα που κάνουμε.Συνήθως χρησιμοποιούμε τον οριζόντιο άξονα για τον χρόνο για να δείξουμε την χρονική συσχέτιση των γεγονότων.Ακόμα έχουμε την δυνατότητα ,για χάρη της καλύτερης παρουσίασης αποτελεσμάτων,να επεξεργαστούμε τα χρώματα στο διάγραμμα και τον τρόπο που εμφανίζονται.



Εικόνα 32: Παραμετροποίηση Dotted Chart

Τα γεγονότα εμφανίζονται πάντα με κουκίδες,ενώ ο χρόνος δείχνει την αφετηρία των γεγονότων καθώς και το τελευταίο γεγονός που εκτελέστηκε.Η κυματομορφή του χρόνου μπορεί να δείξει τις στιγμές όπου οι ενέργειες βρισκόντουσαν στην κορύφωση τους και σε ποία σημεία είχαν ελατωθεί . Μέσα από αυτό βλέπουμε τα συνολικά γεγονότα και ίχνη αυτών,που έχουμε βρει μέχρι αυτήν την στιγμή,στη χρονική στιγμή που εκτελέστηκαν.Σαν μορφή διαγράμματος είναι ικανό να προσφέρει μία αξιόπιστη ματιά στο αρχείο και τα συμβάντα,καθώς και τα μοτίβα που σχηματίζονται κατά την διάρκεια του χρόνου.



Εικόνα 33: Παράδειγμα Dotted Chart

Στα παραπάνω επισημασμένα σημεία βλέπουμε τις χρονικές στιγμές ,όπου υπήρχε μεγάλη εκτέλεση διεργασιών,καθώς έχουμε πυκνά τμήματα από κουκίδες που αναπαρηστούν διαδικασίες.Στα σημεία που οι κουκίδες είναι πιο αραιές έχουμε λιγότερες διαδικασίες,άρα και μικρότερη απόδοση.

3.2.3 PROCESS MODEL

Όπως αναφέραμε και παραπάνω πιο θεωρητικά ,υπάρχουν 4 είδη αλγόριθμων που χρησιμοποιούμε για την δημιουργία του process model από κάθε αρχείο καταγραφής συμβάντων.Στο παράδειγμα μας δουλεύουμε με τους alpha,heuristic,inductive visual και fuzzy miners.Κάθε ένας έχει διαφορετική χρησιμότητα και μπορεί να ωφελήσει σε διάφορες καταστάσεις.Μέσα από την χρήση τους ,είμαστε σε θέση να παράξουμε το μοντέλο διαδικασιών,που δείχνει την ροή εργασιών σε κάθε αρχείο και να δούμε πόσο συμφωνεί με τα δεδομένα που διαθέτουμε.Αξίζει να αναφέρουμε ότι εφαρμόζουμε έναν τέτοιο αλγόριθμο είτε στο αρχικό αρχείο εισόδου ,είτε στο φιλτραρισμένο αρχείο,για να δούμε πως συμπεριφέρεται για ένα συγκεκριμένο γεγονός ή ίχνος.Στη συνέχεια της διπλωματικής θα ασχοληθούμε και με τις δύο περιπτώσεις. [14]

Alpha miner

Ξεκινώντας θα ασχοληθούμε με τον alpha miner καθώς είναι και ο πρώτος τέτοιο αλγόριθμος που φτιάχτηκε.Παρόλα αυτά, ο αλγόριθμος δεν θεωρείται ως πολύ πρακτική τεχνική εξόρυξης καθώς έχει προβλήματα με θόρυβο, σπάνια/ελλιπή συμπεριφορά και πολύπλοκες διαδρομές που δυσκολεύουν την διαδικασία.Ο αλγόριθμος alpha είναι απλός και πολλά από τα χαρακτηριστικά του έχουν ενσωματωθεί σε πιο σύνθετους και ισχυρούς αλγόριθμους,οι οποίοι είναι και αρκετά νεώτεροι.Ο alpha miner παίρνει ένα αρχείο καταγραφής συμβάντων ως είσοδο και παράγει ένα δίκτυο Petri(Petri net). Η πρώτη ενέργεια που κάνει ο alpha miner είναι ότι σαρώνει τα ίχνη για να ταξινομήσει τις σχέσεις μεταξύ των συμβάντων.Ενώ χρησιμοποιεί τις σχέσεις για να δημιουργήσει έναν πίνακα με βήματα αποστάσεων. Υπάρχουν αρκετές παραλλαγές που διαθέτει alpha miner στο Prom Tools και συγκεκριμένα τους alpha+, alpha++, alpha#, alphaR.Ο alpha αλγόριθμος αναγνωρίζει τέσσερις σχέσεις μεταξύ δραστηριοτήτων

1) “>” ή αλλιώς direct(κατευθείαν),όπου ελέγχει για κάθε ζεύγος δραστηριοτήτων άν το A είναι απευθείας ακολουθούμενο από το B σε ένα ίχνος ($A > B$)

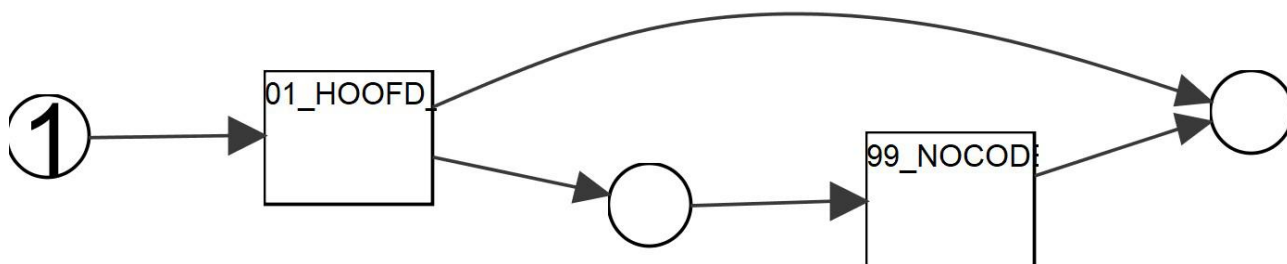
2) “→” ή αλλιώς sequence(ακολουθία),αν το A ακολουθείται απευθείας από το B, αλλά το B ποτέ αμέσως μετά το A ,συμπεραίνουμε ότι το A είναι σε σειρά με το B

3) “||” ή αλλιώς parallel(παράλληλα),αν και το A ακολουθείται απευθείας από το B, καθώς και το B είναι αμέσως μετά το A, συμπεραίνουμε ότι το A και το B είναι παράλληλα. Εκτελούνται δηλαδή και τα δύο

4) “#” ή αλλιώς χωρίς σχέση άμεσα,όταν δηλαδή μετά το A δεν ακολουθεί ποτέ το B και όταν μετά το B δεν ακολουθεί ποτέ το A.Τότε τα δύο γεγονότα δεν έχουν καμία σύνδεση

→ Έξοδος :petri net

→ Πότε να χρησιμοποιείται: Σε περιπτώσεις που θέλουμε να πάρουμε μία πρώτη ματιά από τα δεδομένα, να βγάλουμε κάποια πρώιμα συμπεράσματα και μετά να εφαρμόσουμε εξειδικευμένες τεχνικές. Ο alpha miner είναι απλός από επιστημονικής απόψεως αλλά για τα αρχεία καταγραφής της πραγματικής ζωής, δεν είναι σχεδόν ποτέ η σωστή επιλογή. Δεν θα λειτουργήσει ορθά. Θα δώσει ένα αποτέλεσμα, αλλά δεν θα είναι αξιόπιστο. [15]



Εικόνα 34: Παράδειγμα Alpha miner

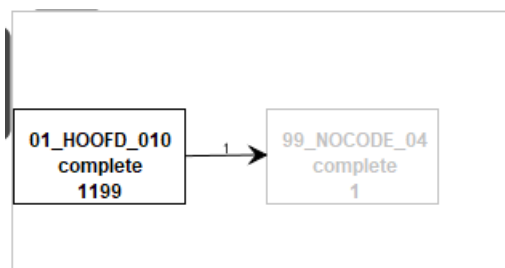
Heuristic miner

Ο heuristic miner είναι στην πράξη μία βελτιωμένη έκδοση του alpha miner που προαναφέραμε. Δημιουργήθηκε λίγο αργότερα και έχει ως λειτουργία να παίρνει τις συχνότητες των γεγονότων και των ακολουθιών, έχοντας αυτά υπόψη κατά την κατασκευή του μοντέλου διαδικασιών. Με αυτόν τον τρόπο μπορεί να φιλτράρει μια θορυβώδη συμπεριφορά ή μια σπάνια συμπεριφορά και να είναι σε θέση να ανιχνεύει σύντομα μονοπάτια, επιτρέποντας την παράλειψη μεμονωμένων ενεργειών. Σε αντίθεση με τον alpha όπου δυσκολεύοταν σε περιπτώσεις θορύβου ή σπάνιας συμπεριφοράς. Η σκέψη πίσω από αυτό είναι η αποκοπή σπάνιων μονοπατιών από το διάγραμμα δραστηριοτήτων.

Ο heuristic αναγνωρίζει τέσσερα είδη σχέσεων, όπως και ο alpha, αλλά ενσωματώνει και την συχνότητα εμφάνισης του κάθε ίχνους.

→ Έξοδος: Petrinet

→ Πότε να χρησιμοποιείται: Όταν έχετε πραγματικά δεδομένα με όχι πάρα πολλά διαφορετικά συμβάντα ή όταν χρειάζεστε ένα μοντέλο Petri net για περαιτέρω ανάλυση στο ProM [16]



Εικόνα 35: Παράδειγμα Heuristic miner

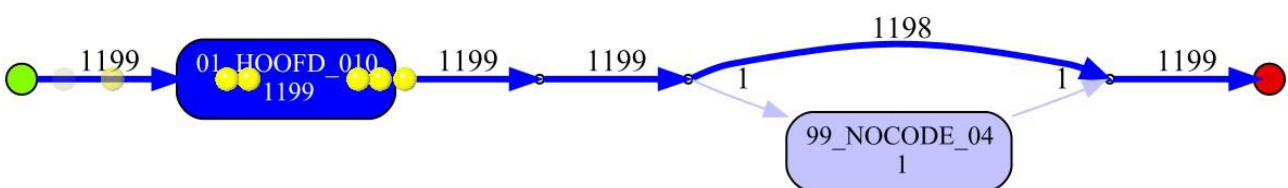
Inductive Visual miner

Όσον αφορά τον inductive αλγόριθμο έχουμε να κάνουμε με τον πιο αποδοτικό εκ των τεσσάρων, ίσως και την καλύτερη επιλογή που έχουμε, όταν θέλουμε να δημιουργήσουμε το μοντέλο διαδικασιών από ένα αρχείο καταγραφής συμβάντων. Εφαρμόζει σε μεγάλο βαθμό την τεχνική divide and conquer (διαίρε και βασίλευε), όπου χωρίζει το αρχείο, κατασκευάζει μέρη του ολικού μοντέλου διαδικασιών και επεξεργάζεται το καθένα ξεχωριστά. Αυτό του δίνει ταχύτητα και ευελιξία στις δοκιμές, πράγμα που λείπει από τους υπόλοιπους. Βέβαια εξαρτάται πάντα και από το τελικό σκοπό που θέλουμε να επιτύχουμε αλλά σε γενικές γραμμές προσφέρει τεράστια πλεονεκτήματα όπως:

- το μοντέλο είναι πάντα σε θέση να αναπαραστήσει επακριβώς τις διαδικασίες του αρχείου καταγραφής συμβάντων
- έχει την δυνατότητα να χειριστεί περιπτώσεις θορύβου, σπάνιας συμπεριφοράς και περιπτώσεων με μοτίβα χαμηλής συχνότητας εμφάνισης
- είναι αρκετά παραμετροποιήσιμο και επιτρέπει στο χρήστη να κάνει αρκετούς πειραματισμούς ανάλογα με τους σκοπούς του
- προσφέρει ζωντανή παρουσίαση του μοντέλου σε πραγματικό χρόνο, δίνοντας την ευκαιρία να παρατηρήσουμε τι ακριβώς συμβαίνει και ποια χρονική στιγμή
- η έξοδος που παράγει, συνήθως αρχείο, μπορεί εύκολα να επεξεργαστεί και να φιλτραριστεί χωρίς κάποια μετατροπή, δίνοντας υψηλή μεταφερσιμότητα στα αρχεία

→ Έξοδος: Petrinet

→ Πότε να χρησιμοποιείται: Σε προβλήματα όπου θέλουμε να αντιμετωπίσουμε θορυβώδη συμπεριφορά από δεδομένα. Όταν θέλουμε να βρούμε τον πιο προφανή διαχωρισμό στο αρχείο συμβάντων, να αναγνωρίσουμε την πράξη και να επεξεργαστούμε το κάθε ένα χωριστά.



Εικόνα 36: Παράδειγμα Inductive Visual miner

Fuzzy miner

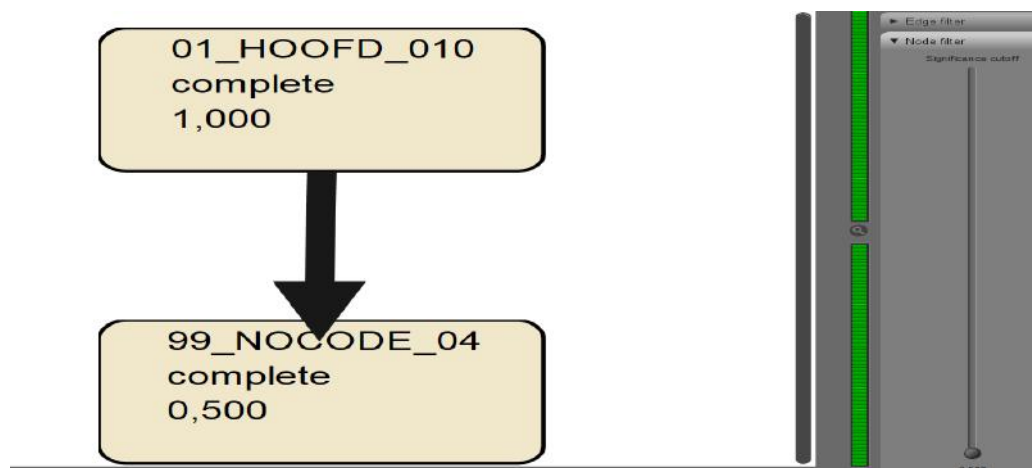
Ο fuzzy miner αποτελεί μία διαφοροποίηση των προηγούμενων αλγορίθμων, καθώς δεν λειτουργεί με τον παραδοσιακό τρόπο. Αντίθετα με τους υπόλοιπους αλγόριθμους δημιουργίας μοντέλου διαδικασιών, δεν διαχειρίζεται petri nets αλλά γραφήματα. Σε αυτήν την περίπτωση στην δημιουργία

του μοντέλου,παίζουν σημαντικό ρόλο η σημασία των διεργασιών και οι σχέσεις μεταξύ αυτών.Ακόμα σε αντίθεση με τους άλλους αλγόριθμους ,προσφέρεται η δυνατότητα αφαίρεσης και επεξεργασίας των στοιχείων του αρχείου,με αποτέλεσμα οι διεργασίες και οι σχέσεις τους να επεξεργάζονται ή να αφαιρούνται ομαδικά.

Η σπουδαιότητα αυτού του αλγόριθμου προκύπτει από την ανάγκη ύπαρξης τεχνικών διαχείρισης ασαφών και σπάνιων γεγονότων.Όλοι οι αλγόριθμοι μέχρι τότε υστερούσαν σε αυτό το κομμάτι, μέχρι την στιγμή που δημιουργήθηκε ο Fuzzy miner.Δίνει την δυνατότητα διεκπεραίωσης τέτοιων καταστάσεων με ασταθή δεδομένα και επίλυσης αυτών των ζητημάτων.[17]

→ Έξοδος:FuzzyModel

→ Πότε να χρησιμοποιείται: Όταν έχετε πολύπλοκα και μη δομημένα δεδομένα καταγραφής ή όταν θέλετε να απλοποιήσετε το μοντέλο με διαδραστικό τρόπο



Εικόνα 37: Παράδειγμα Fuzzy miner

3.2.4 FILTERING EVENT LOG

Το φιλτράρισμα του αρχείου καταγραφής συμβάντων αποτελεί ίσως το σημαντικότερο βήμα στο process mining,καθώς με αυτόν τον τρόπο εισερχόμαστε στο πρακτικό κομμάτι αυτής της τεχνικής.Τα δεδομένα που περιλαμβάνει ένα αρχείο ,θα πρέπει να προέρχονται από διαφορετικές πηγές για να καλύψουμε μία πληθώρα στοιχείων από διαφορετικές περιπτώσεις και μοτίβα.Στο δικό μας παράδειγμα τα αρχεία των δήμων θα πρέπει να περιλαμβάνουν ενέργειες από πολλούς εργαζομένους,διαφορετικούς τομείς,εποχές και περιοχές.Με βασικό σκοπό να διασφαλίσουμε μία ποικιλομορφία των δεδομένων,για να είναι όσο το δυνατόν πιο αντιπροσωπευτικά για το

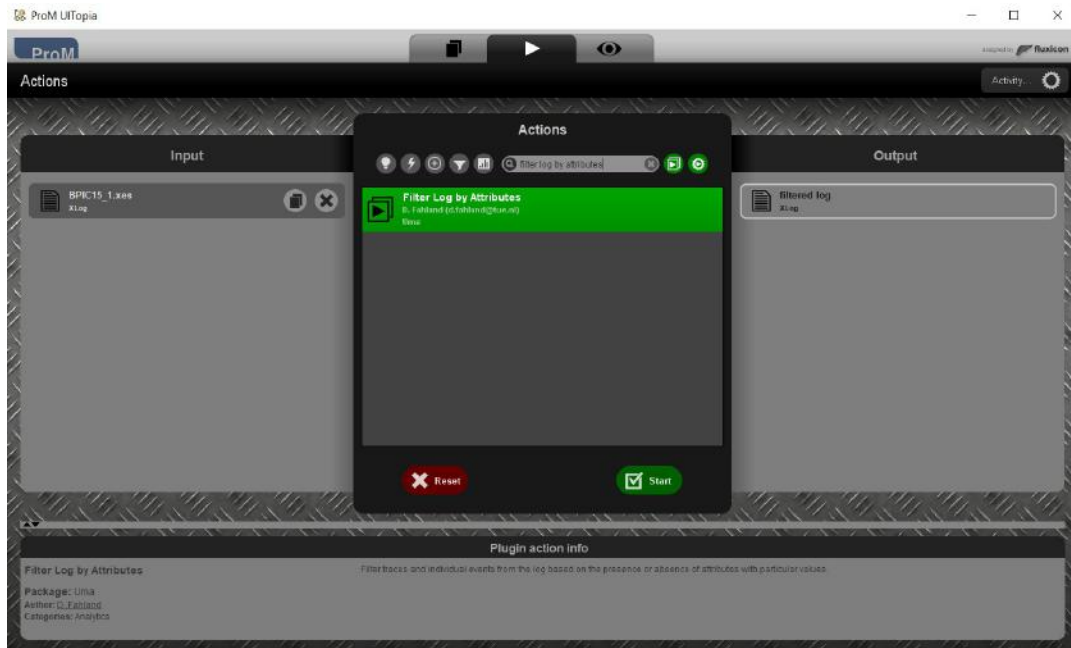
μεγαλύτερο μέρος των εργασιών. Αυτό χρειάζεται διότι μπορεί να υπάρχουν λάθη ή ακόμα οι μετρήσεις να μην περιλαμβάνουν μεγάλο αριθμό δειγμάτων. Κύριως στόχος αποτελεί το φιλτράρισμα των αρχείων καταγραφής συμβάντων που περιλαμβάνουν λάθη και πρέπει να διορθωθούν. Ακόμα είναι χρήσιμο να γίνει σε περιπτώσεις ασυνήθιστων συμπεριφορών, όπως πολύ μεγάλες ή μικρές τιμές ή κάποια άλλη ανομοιογένεια που έχουμε παρατηρήσει και θέλουμε να αναλύσουμε.

Με αυτά τα κριτήρια συνήθως εφαρμόζουμε φίλτρα στα αρχεία καταγραφής συμβάντων είτε για να τα καθαρίσουμε από λάθη ή άκυρες τιμές, είτε για να περιορίσουμε τα δεδομένα σε μία μικρότερη κλίμακα, την οποία αργότερα θα μελετήσουμε την συμπεριφορά της, τα μοτίβα και το μοντέλο διαδικασιών. Με την πλατφόρμα του Prom Tools στις περισσότερες υποθέσεις, φιλτράρουμε για να αφαιρέσουμε κάποια ίχνη, να αφαιρέσουμε ή προσθέσουμε κάποια γεγονότα και να επεξεργαστούμε ή παραμετροποιήσουμε κάποια από αυτά.

Εν συνεχεία το επόμενο βήμα θα ήταν να φιλτράρουμε το αρχείο εισόδου με κάποιο από τα διαθέσιμα plug in που διαθέτει το Prom Tools. Μερικά από αυτά είναι το filter log by attributes, filter events based on attribute values και filter log on events attributes values. Με κάποιο από τα παραπάνω αξιόπιστα plug in μπορούμε να επεξεργαστούμε το αρχείο με βάση μία ιδιαιτερότητα (ή ανομοιογένεια) που συναντήσαμε πιο πριν, όπως την πιο συνηθισμένη ή σπάνια δραστηριότητα, ποιος δήμος εκτελεί τις περισσότερες εργασίες ή ποιος τις λιγότερες. Επισημαίνεται ότι στο filter log by attributes υπάρχει και η επιλογή trace length για να δούμε ποιες και πόσες υποθέσεις σε κάθε δήμο, αποτελούνται από πολλές δραστηριότητες. Στην επιλογή input τοποθετούμε το αρχείο εισόδου, στην καρτέλα actions το επιθυμητό plug in και τέλος η καρτέλα output εμπεριέχει το τελικό αποτέλεσμα. [18]

Έτσι λοιπόν στην καρτέλα Actions έχουμε την επιλογή αρχείου εισόδου και επιθυμητού plug in Actions





Εικόνα 38: Καρτέλα Actions

Έτσι με αυτήν την τεχνική εμφανίζονται τα αποτελέσματα μόνο για το εκάστοτε ζήτημα που θέλουμε να ψάξουμε και μπορούμε να δούμε μεμονωμένα πως συμπεριφέρεται. Για παράδειγμα μπορούμε να ψάξουμε ένα συγκεκριμένο event ή οποιοδήποτε attribute μας ενδιαφέρει.

Έπειτα αφού έχουμε το φιλτραρισμένο αρχείο με το δικό μας κριτήριο, μπορούμε να εφαρμόσουμε ξανά έναν miner που διαθέτει το ProM Tools. Στην διπλωματική ασχολούμαστε περισσότερο με τους Alpha, Heuristic, Inductive Visual και Fuzzy miner για να παραχθούν τα αντίστοιχα μοντέλα διαδικασιών (process model) και να δούμε ποίο από αυτά εφαρμόζεται καλύτερα και συμφωνεί περισσότερο με το σετ δεδομένων που διαθέτουμε. Μπορούμε τέλος να εφαρμόσουμε κάποια τεχνική εικονικοποίησης στο φιλτραρισμένο αρχείου εξόδου, όπως explore event log ή dotted chart, για να παρατηρήσουμε τα νέα στοιχεία.

ΚΕΦΑΛΑΙΟ 4

4.1 ΜΕΛΕΤΗ ΤΗΣ ΔΙΠΛΩΜΑΤΙΚΗΣ

Στα προηγούμενα μέρη αναφέραμε ότι το πλήθος των εργασιών που λαμβάνουν χώρα καθημερινά είναι τεράστιος. Τόσο σε ατομικό επίπεδο όσο και σε συλλογικό, συνέχεια και ακατάπαυστα εκτελούνται ενέργειες, είτε έχουν να κάνουν με επαγγελματικές υποχρεώσεις είτε με προσωπικά ενδιαφέροντα του καθενός. Οι διαδικασίες συμβαίνουν καθόλη την διάρκεια της ζωής μας και οτιδήποτε κάνουμε, θεωρείται μία ενέργεια. Από το πρωινό ξύπνημα μέχρι και τις ενέργειες που κάνουμε στον εργασιακό χώρο, όλα αυτά είναι διαδικασίες. Όπως γίνεται αντιληπτό κάθε άτομο αποτελείται από εκατοντάδες εργασίες σε μία μόνο μέρα.

Ειδικά στον εργασιακό τομέα που μας ενδιαφέρει περισσότερο, οι εργασίες είναι πολλές και φέρουν πολλά δεδομένα η καθεμία, τα οποία πρέπει να διαχειριστούν και να αναλυθούν. Μία επιχείρηση ή ένας οργανισμός οφείλει να γνωρίζει τις διαδικασίες που τρέχουν καθημερινά, να είναι ενήμερος για την απόδοση τους και τους τρόπους που μπορούν να συντελέσουν στην βελτίωση αυτής. Ειδικά στον κλάδο της τοπικής αυτοδιοίκησης με τον οποίο ασχοληθήκαμε στην διπλωματική, πρέπει να γίνεται προσπάθεια βελτίωσης των εργασιών και συνεπώς των υπηρεσιών που προσφέρονται στους πολίτες της χώρας. Υπάρχουν πολλά αρχεία δεδομένων, με στοιχεία που συλλέγονται από δήμους με επισημασμένα δεδομένα για τη δραστηριότητα και τις πληροφορίες της χρονικής σφραγίδας που μπορούν να χρησιμοποιηθούν ως αρχεία καταγραφής συμβάντων. Χρησιμοποιώντας τεχνικές εξόρυξης διεργασιών, μπορούμε να ανακαλύψουμε το μοντέλο διαδικασιών, να αναλύσουμε τις εργασίες και το πως λειτουργούν τα άτομα που τις εκτελούν. Αναλύοντας αυτά τα δεδομένα, μπορούμε να βγάλουμε συμπεράσματα για αυτούς. Μπορούμε να μάθουμε για κάθε δραστηριότητα πόσες φορές εκτελείται και από ποιους υπαλλήλους, αν εκτελείται συχνά ή σπάνια. Ακόμα βρίσκουμε τους δήμους που εκτελούν τις περισσότερες εργασίες σε συλλογικό επίπεδο και ποιοι δήμοι έχουν τους πιο αποδοτικούς υπαλλήλους. Από την άλλη πλευρά σε περιπτώσεις χαμηλής απόδοσης, είμαστε σε θέση να βρούμε σε ποία σημεία παρατηρείται bottleneck της απόδοσης και με τα κατάλληλα εργαλεία να το αντιμετωπίσουμε. Έχουμε την ικανότητα να δημιουργούμε μοτίβα, με βάση τα συμπεράσματα που εξάγουμε, ώστε να μπορούμε να προβλέψουμε τα ευαίσθητα σημεία πριν συμβούν και να είμαστε μπροστά από τις εξελίξεις και όχι πίσω από αυτές.

Οι εργασίες που εκτελούνται καθημερινά έχουν πολλές παραμέτρους και ποικίλλουν από διεργασία σε διεργασία. Σε πολλές περιπτώσεις έχουμε να αντιμετωπίσουμε χιλιάδες εργασίες σε πολλούς διαφορετικούς τομείς. Αποτελούνται από αρκετές ιδιότητες, κάθεμια περιγράφει και ένα διαφορετικό πεδίο, ενώ σε πολλά παραδείγματα έχουμε να κάνουμε με εργασίες που αλλάζουν συνεχώς και θα πρέπει να είμαστε σε θέση να τις μελετήσουμε.

Στην μελέτη που έγινε χρησιμοποιήσαμε αρχεία καταγραφής συμβάντων από 5 Ολλανδικούς δήμους στο διάστημα από το 2010 έως το 2015. Τα αρχεία υποστηρίζουν το πρότυπο XES που προαναφέραμε και συγκεντρώθηκαν από τους δήμους και από τα αρμόδια όργανα. Τα συμβάντα αφορούν διαδικασίες/εργασίες που εκτελέστηκαν σε βάθος 5ετίας από τους δήμους και παρουσιάζουν σε καθημερινό επίπεδο τις ενέργειες που έλαβαν χώρα. Κάθε εργασία είχε διαφορετικά χαρακτηριστικά, παρουσιάζοντας ποικίλες διαδικασίες στο αρχείο όπως εργασίες

κατασκευής ή συντήρησης χώρων και εγκαταστάσεων, διοικητικές εργασίες και εργασίες για την εξυπηρέτηση των αναγκών που υπήρχαν. Σε αυτές τις περιπτώσεις βλέπουμε την διαφορετική συμπεριφορά που υπάρχει σε κάθε διαδικασία και δήμο, καθώς και τα μοτίβα που αναπτύσσονται καθόλη την διάρκεια της 5ετίας.

Σκοπός αυτής της διπλωματικής είναι να χρησιμοποιήσει τεχνικές εξόρυξης διεργασιών, για να αναλύσει δυνατότητες και πληροφορίες που μπορούν να αποκαλυφθούν από το αρχείο καταγραφής συμβάντων. Έχουμε ως στόχο να μελετήσουμε και να περιγράψουμε τις διαδικασίες που εκτελούνται, να ερευνήσουμε τις συμπεριφορές και τα μοτίβα που συναντάμε, να εντοπίσουμε παθογένειες ή ανομοιογένειες στα δεδομένα και με την χρήση των κατάλληλων εργαλείων να τις επιλύσουμε.

Θα χρησιμοποιήσουμε το ProM Tools καθώς είναι πολύ ισχυρό ανοιχτού κώδικα λογισμικό, είναι επεκτάσιμο και χρησιμοποιείται σε ερευνητικές εργασίες. Θα ανακαλύψουμε τις διαδικασίες με διάφορους αλγόριθμους και θα συγκρίνουμε τα αποτελέσματα. Επιδίωξη μας είναι, μέσα από την έρευνα να δώσουμε απαντήσεις σε μείζων ερωτήματα, που προβληματίζουν τον τομέα των διαδικασιών

Τα βήματα που θα χρειαστεί να ακολουθήσουμε σε κάθε αρχείο καταγραφής συμβάντων για να πετύχουμε το σκοπό μας είναι τα εξής:

- Import event logs
- Event log visualization
- Discover process models
- Filter event logs

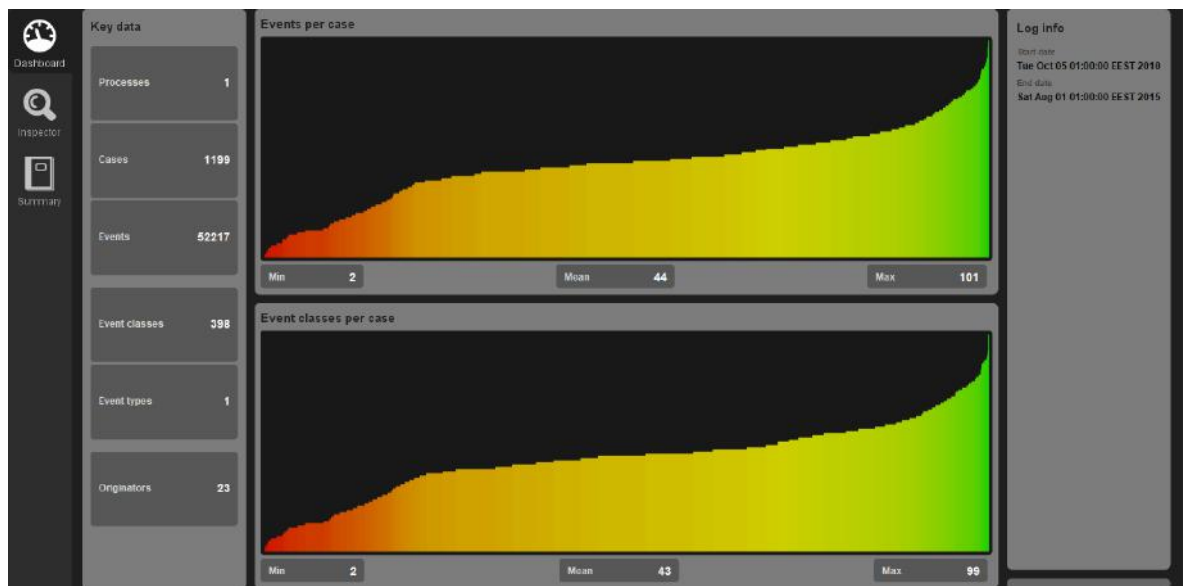
4.1.1 ΕΙΣΑΓΩΓΗ ΑΡΧΕΙΩΝ ΚΑΤΑΓΡΑΦΗΣ ΣΥΜΒΑΝΤΩΝ

Το πρώτο βήμα είναι να εισάγουμε μέσω της επιλογής import, τα αρχεία των δήμων που θα πρέπει να μελετήσουμε και να επεξεργαστούμε. Πρόκειται για αρχεία με πρότυπο xml. Τα αρχεία εμφανίζονται με ονόματα

- DATA1.xml (1ος δήμος)
- DATA2.xml (2ος δήμος)
- DATA3.xml (3ος δήμος)
- DATA4.xml (4ος δήμος)
- DATA5.xml (5ος δήμος)

Παρακάτω φαίνονται οι πίνακες dashboard/log visualizer που έχουμε αναφέρει και πρωτίστως,

1ος δήμος

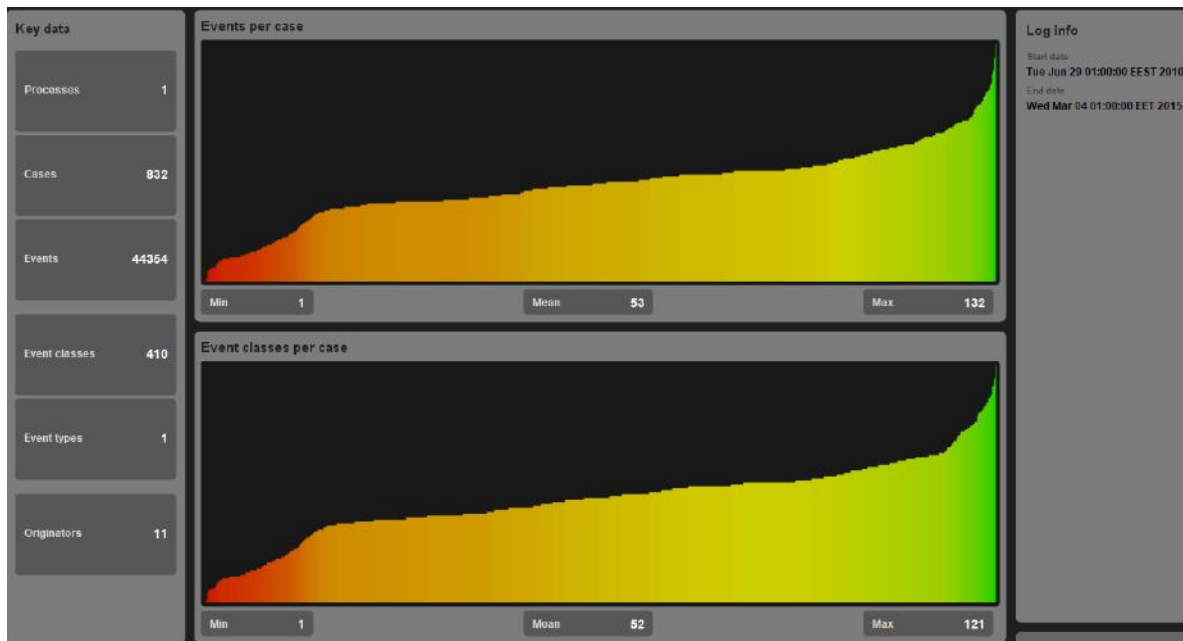


Εικόνα 39: Dashboard 1ου δήμου

Από την παραπάνω εικόνα παίρνουμε τις εξής πληροφορίες:

Processes	1
Cases	1199
Events	52217
Event classes	398
Event types	1
Originators	23
Start date	05/10/2010 01:00:00
End date	01/08/2015 01:00:00

2ος δήμος

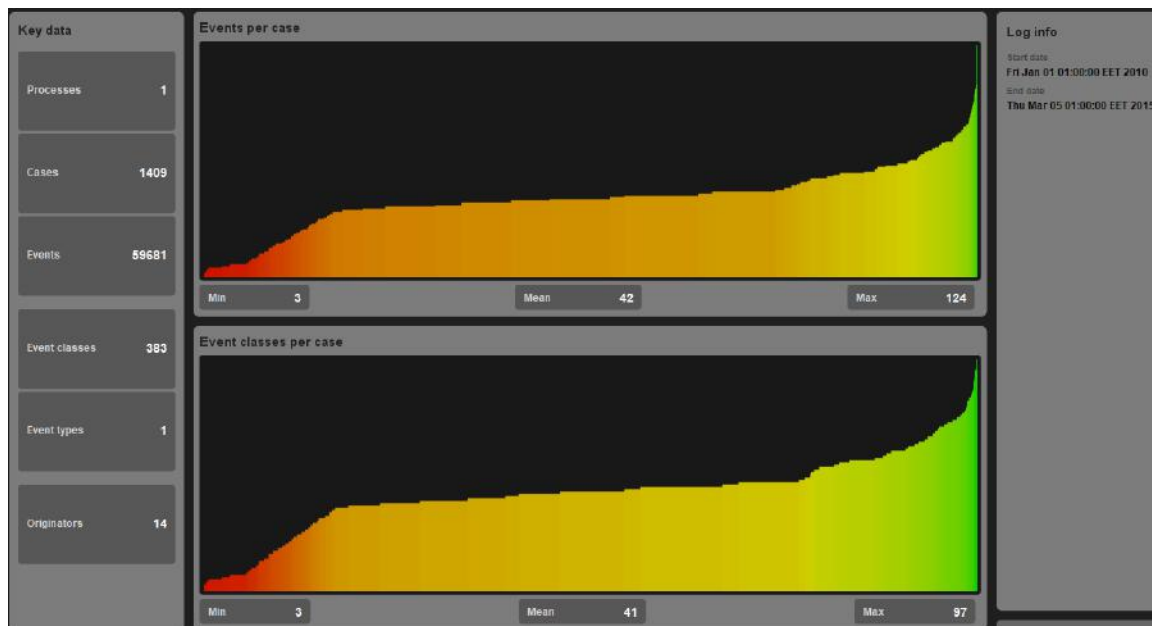


Εικόνα 40: Dashboard 2ου δήμου

Από την παραπάνω εικόνα παίρνουμε τις εξής πληροφορίες:

Processes	1
Cases	832
Events	44354
Event classes	410
Event types	1
Originators	11
Start date	29/06/2010 01:00:00
End date	04/03/2015 01:00:00

3ος δήμος

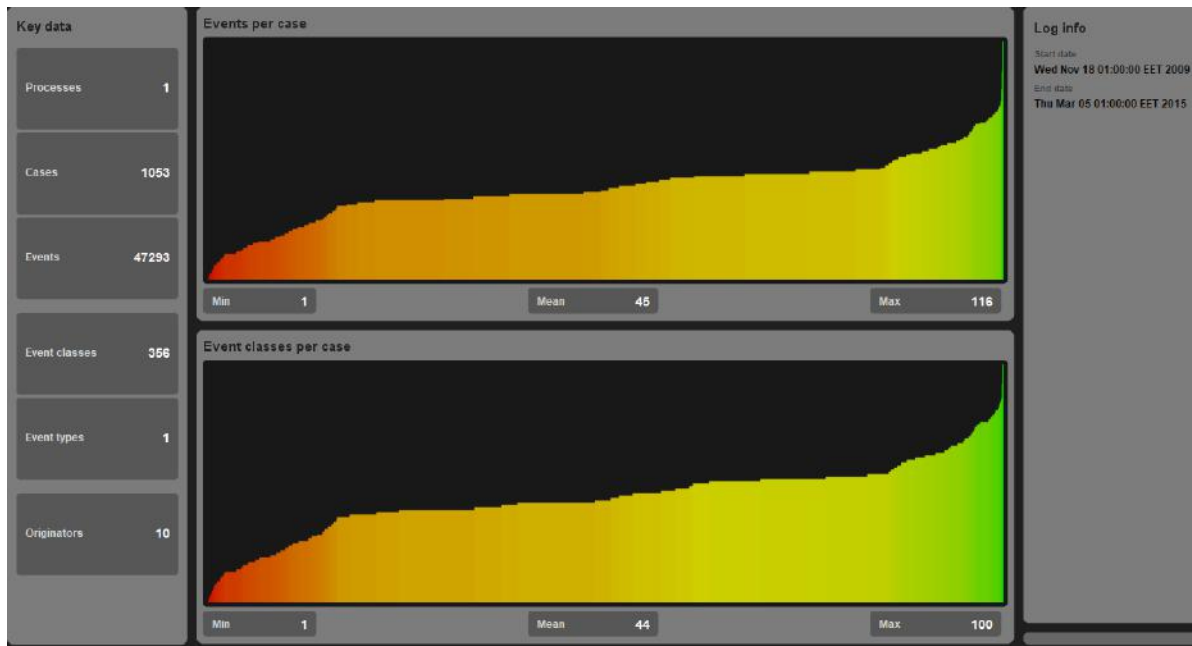


Εικόνα 41: Dashboard 3ου δήμου

Από την παραπάνω εικόνα παίρνουμε τις εξής πληροφορίες:

Processes	1
Cases	1409
Events	59681
Event classes	383
Event types	1
Originators	14
Start date	01/06/2010 01:00:00
End date	05/03/2015 01:00:00

4ος δήμος

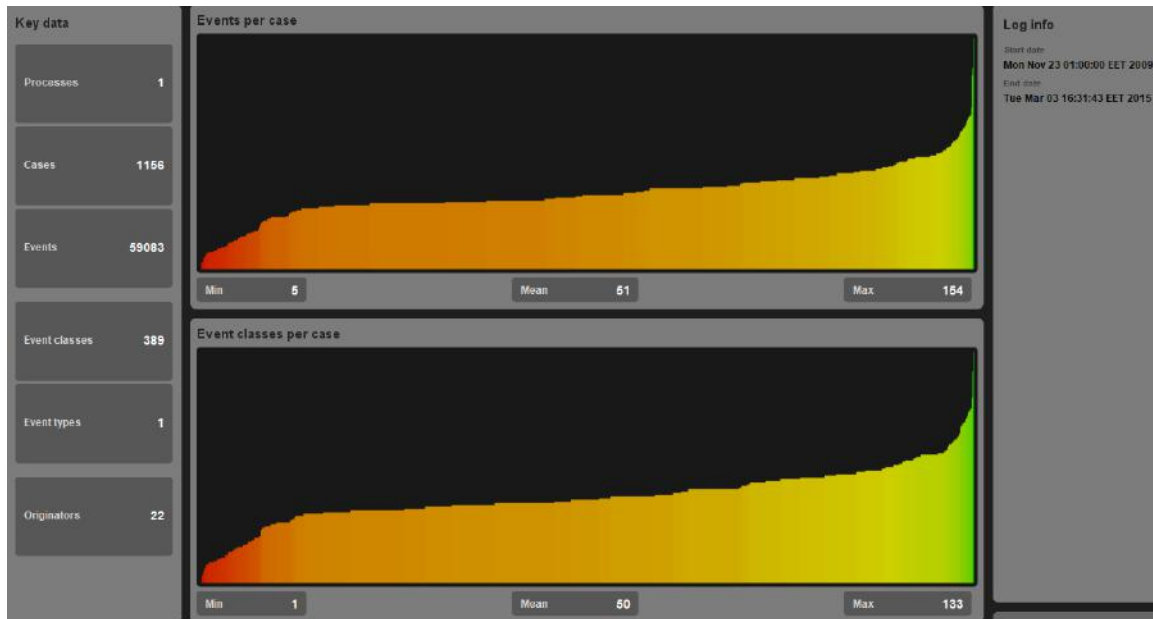


Εικόνα 42: Dashboard 4ου δήμου

Από την παραπάνω εικόνα παίρνουμε τις εξής πληροφορίες:

Processes	1
Cases	1053
Events	47293
Event classes	356
Event types	1
Originators	10
Start date	18/11/2009 01:00:00
End date	05/03/2015 01:00:00

5ος δήμος



Εικόνα 43: Dashboard 5ου δήμου

Από την παραπάνω εικόνα παίρνουμε τις εξής πληροφορίες:

Processes	1
Cases	1156
Events	59083
Event classes	389
Event types	1
Originators	22
Start date	23/11/2009 01:00:00
End date	03/03/2015 01:00:00

Από τις μετρήσεις βγάζουμε κάποια συμπεράσματα για την συμπεριφορά των δήμων και τις διαδικασίες που εκτελούνται όλο αυτό το διάστημα. Πιο συγκεκριμένα έχουμε τον 1ο, 3ο και 5ο δήμο με αρκετές παραπάνω διαδικασίες σε βάθος 5ετίας. Και οι τρεις δήμοι βρίσκονται με πάνω από 50.000 διαδικασίες και οι υποθέσεις τους ξεπερνούν τις 1.000.0 1ος και ο 5ος δήμος διαθέτουν αρκετούς υπαλλήλους, ενώ λιγότερους έχει ο 3ος, πράγμα που σημαίνει ότι εκτελεί αρκετές διαδικασίες με λιγότερους εργαζομένους. Ακόμα λιγότερους έχει ο 2ος και 4ος δήμος, αντίστοιχα με τις λιγότερες διαδικασίες που εκτελούν. Όσον αφορά τις ημερομηνίες έναρξης και λήξης των διαδικασιών στα αρχεία καταγραφής συμβάντων, ο ακόλουθος πίνακας παρουσιάζει συγκεντρωμένα τα στοιχεία.

ΔΗΜΟΣ
1ος

ΑΡΧΙΚΗ ΗΜΕΡΟΜΗΝΙΑ
5/10/2010

ΤΕΛΙΚΗ ΗΜΕΡΟΜΗΝΙΑ
1/08/2015

2ος	29/06/2010	4/03/2015
3ος	1/01/2010	5/03/2015
4ος	18/11/2009	5/03/2015
5ος	23/11/2009	3/03/2015

Ανακαλύπτουμε ότι όλα τα αρχεία δεδομένων ολοκληρώνονται το 2015,συγκεκριμένα τα 4 από τα 5 τον Μάρτιο,ενώ το τελευταίο τον Αύγουστο του ίδιου έτους.Ακόμα βλέπουμε ότι ο 4ος και 5ος δήμος ξεκινάνε την καταγραφή στα αρχεία νωρίτερα,συγκεκριμένα τέλη του 2009,σε αντίθεση με τους υπόλοιπους που αρχίζουν μέσα στο 2010.Ειδικότερα ο 1ος δήμος ξεκινάει σχεδόν 1 χρόνο μετά από τους πρώτους,την καταγραφή των log files και ο 2ος δήμος περίπου 7 μήνες μετά.Εντυπωσιακό αν αναλογιστούμε την υψηλή απόδοση και παραγωγικότητα του 2ου δήμου,σε σχέση με το χρονικό διάστημα που είχε στην διάθεση του.

Στη συνέχεια των προηγούμενων μετρήσεων υπολόγισα και έναν μέσο όρο διεργασιών ανά ημέρα για κάθε δήμο,γνωρίζοντας τον αριθμό των εργασιών και τις διαθέσιμες μέρες που είχε για να τις εκτελέσει.

Καταλήγουμε στα στατιστικά:

1ος δήμος → 7,25 events ανά ημέρα
 2ος δήμος → 9,22 events ανά ημέρα
 3ος δήμος → 7,80 events ανά ημέρα
 4ος δήμος → 7,42 events ανά ημέρα
 5ος δήμος → 9,06 events ανά ημέρα

Φτάνοντας στο συμπέρασμα ότι ο 2ος και 5ος δήμος είχαν την υψηλότερη απόδοση,καθώς τα αποτελέσματα δείχνουν ότι εκτελούνταν πάνω από 9 εργασίες καθημερινώς,σε αντίθεση με τους υπόλοιπους 3 δήμους που κυμαίνονται στις 7 εργασίες ανά ημέρα.

Πριν περάσουμε στις μετρήσεις με τα εργαλεία του Prom,επειδή θα συναντήσουμε πολλές διεργασίες με τους κωδικούς τους και θα πρέπει να είναι κατανοητές απολύτως,παραθέτω έναν πίνακα με τις πιο δημοφιλείς διεργασίες με την ονομασία της στα Αγγλικά δίπλα από το κωδικό που έχουν ήδη.Σκοπός είναι να μας βοηθήσει να καταλάβουμε λίγο καλύτερα τις διαδικασίες που λαμβάνουν χώρα και να μπορούμε να τις διαχειριστούμε καλύτερα,όχι μόνο με έναν τυχαίο κωδικό που τους ανατέθηκε.

Έτσι έχουμε τις εξής:

EVENT CODE	EVENT NAME (EN)
01_HOOFD_010	register submission date request
01_HOOFD_011	OLO messaging active
01_HOOFD_015	phase application received
01_HOOFD_020	send confirmation receipt
01_HOOFD_030_	enter senddate acknowledgement
01_HOOFD_050	applicant is stakeholder

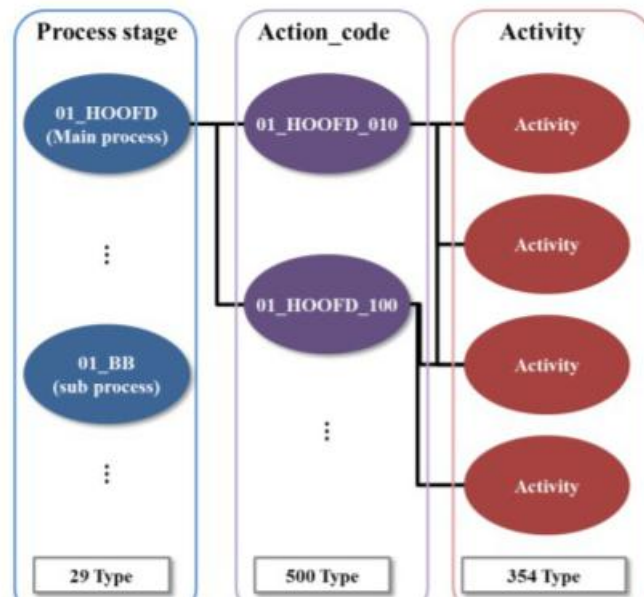
01_HOOFD_061	start WABO procedure
01_HOOFD_065	create procedure confirmation
01_HOOFD_040	forward to the competent authority
01_HOOFD_060	regular procedure without MER
01_HOOFD_065	send procedure confirmation
01_HOOFD_090	publish
01_HOOFD_101	registration date publication
05_EIND_010	terminate on request
16_LGSV_010	calculate provisional charges
06_VD_010	extend procedure term
01_HOOFD_540	suspend term
01_HOOFD_110_	create subcases completeness
01_HOOFD_110	treat subcases completeness
01_HOOFD_180	procedure change
15_NGV_010	request further information
01_HOOFD_195	phase application receptive
01_HOOFD_196	procedure change after completeness
01_HOOFD_200	send letter in progress
01_HOOFD_250_	create subcases content
01_HOOFD_250	treat subcases content
01_HOOFD_250_	completed subcases content
01_HOOFD_190	request complete
09_AH_I_010	article 34 WABO applies
01_HOOFD_370	assessment of content completed
01_HOOFD_265	phase advise known
01_HOOFD_380	grounds for refusal
13_CRD_010	coordination of application

Όπως παρατηρούμε και παραπάνω οι διεργασίες με κωδικό HOOFD είναι οι πιο συνηθισμένες στους 5 δήμους, ενώ πολλές εξ αυτών υπάρχουν και στους 5 δήμους. Για παράδειγμα οι δήμοι έχουν κοινές τις διεργασίες με κωδικό HOOFD

- send confirmation receipt με κωδικό 01_HOOFD_020
- procedure change με κωδικό 01_HOOFD_180
- enter senddate decision environmental permit με κωδικό 01_HOOFD_510
- register submission date request με κωδικό 01_HOOFD_010
- phase application received με κωδικό 01_HOOFD_015

Και επίσης τη διεργασία με κωδικό AWB45.[19]

Ακόμα πρέπει να αναφέρουμε ότι οι κωδικοί των εργασιών φέρουν και κάποια άλλη σημασιολογία. Το πρώτο μέρος του κωδικού '01_HOOFD_xxx' σηματοδοτεί το τομέα στον οποίο εντάσσεται η εργασία, ενώ το πίσω μέρος 'xx_xxxx_001' υποδηλώνει τη σειρά με την οποία εκτελούνται οι δραστηριότητες, όπου το πρώτο ψηφίο υποδηλώνει συχνά μια φάση μέσα σε μια διαδικασία. Στο παράδειγμα της διπλωματικής μας έχουμε 29 διαφορετικά είδη εργασιών, 500 διαφορετικούς κωδικούς για τις εργασίες και 354 τύπους δραστηριοτήτων.



Εικόνα 44: Ανάλυση κωδικών διαδικασιών

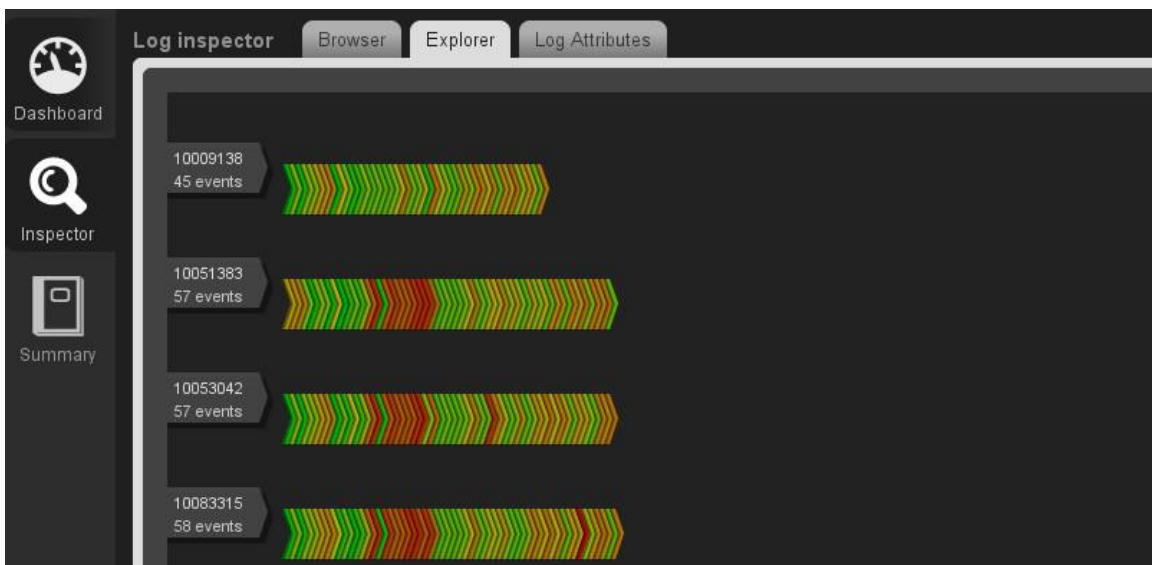
4.1.2 EVENT LOG VISUALIZATION

INSPECTOR/EXPLORER

Στην συνέχεια με το επόμενο παράθυρο Inspector μπορούμε να δούμε τις υποθέσεις του κάθε αρχείου, από ποιές διαδικασίες αποτελούνται και τα attributes που διαθέτουν. Έπειτα με την χρήση του explorer έχουμε την δυνατότητα να παρατηρήσουμε τις υποθέσεις και τον αριθμό των διαδικασιών που περιλαμβάνει η κάθε μία. Ακόμα όπως προαναφέραμε μπορούμε για κάθε διαδικασία να δούμε στοιχεία για αυτήν και με βάση το χρώμα που έχει, πόσο συχνά εμφανίζεται.



Εικόνα 45: Καρτέλα browser του inspector για τους δήμους της μελέτης



Εικόνα 46: Καρτέλα explorer του inspector για τους δήμους της μελέτης

Στα παραπάνω παραδείγματα παρατηρούμε μέσω του Inspector όλα τα ίχνη με τις διαδικασίες τους και στην δεξιά πλευρά τα attributes του κάθε ίχνους. Από το παράδειγμα βλέπουμε τα εξής attributes για την υπόθεση με κωδικό 10009138 με 45 διαδικασίες, όπου πρώτη είναι η ενέργεια με κωδικό 01_HOOFD_010

CONTINUOUS TYPED:

→ SUMIpleges (πληρωθέν ποσό για την αίτηση άδειας)

LITERAL TYPED:

- IdofConceptCase (αναγνωριστικό υπόθεσης)
- Responsible actor (υπεύθυνος οργανισμός/υπάλληλος)
- caseStatus (κατάσταση υπόθεσης ανοικτή ή κλειστή)
- caseType (ένας αριθμός ίδιος για όλες τις υποθέσεις στο αρχείο)

- conceptname (αναγνωριστικό διεργασίας)
- last_phase (η τελευταία διαδικασία στο αρχείο)
- parts (περιγραφή για τον τύπο της διεργασίας π.χ. χτίσιμο, συντήρηση)
- requestComplete (ολοκλήρωση ή όχι της υπόθεσης)
- termName (πλήρη ονομασία της υπόθεσης)

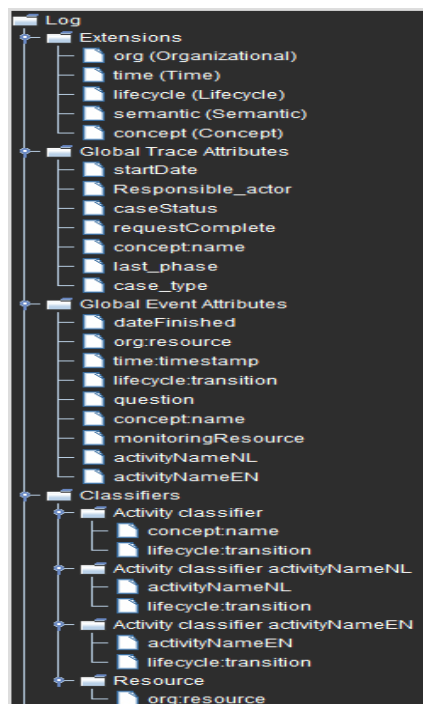
TIMESTAMP TYPED:

- endDate (ημερομηνία ολοκλήρωσης)
- endDatePlanned (προγραμματισμένη ημερομηνία ολοκλήρωσης)
- startDate (ημερομηνία έναρξης)

Ενώ στην συνέχεια στον Explorer βλέπουμε για τα cases, τα event που περιέχουν και πόσο συχνά εμφανίζονται. Για παράδειγμα στο προαναφερθέν ίχνος 10009138 βλέπουμε ενέργειες με ανοιχτό πράσινο χρώμα έως κίτρινο, που σημαίνει ότι αποτελείται από σχετικά δημοφιλή event. Ενώ αντίθετα το επόμενο ίχνος 10051383 περιλαμβάνει ενέργειες με κόκκινο χρώμα, δηλαδή με μικρό αριθμό εκτέλεσης.

LOG ATTRIBUTES

Το επόμενο παράθυρο μας δίνει πληροφορίες για τα extensions (επεκτάσεις), global trace attributes, global event attributes και classifiers (διαχωριστές) που έχουν χρησιμοποιηθεί. Μας δείχνει τον τρόπο που έχουν δομηθεί τα δεδομένα στα αρχεία.



Εικόνα 47: Καρτέλα Log Attributes για τη δομή των αρχείων της μελέτης

- Επεκτάσεις για την δομή, τον χρόνο, τον κύκλο ζωής, την σημασιολογία και την υπόθεση του κάθε ίχνους και ενέργειας που υπάρχει στο αρχείο καταγραφής συμβάντων

→ χαρακτηριστικά για το κάθε ίχνος:ημερομηνία έναρξης,υπεύθυνος υπάλληλος, κατάσταση υπόθεσης ανοικτή ή κλειστή,ολοκλήρωση ή όχι της υπόθεσης, αναγνωριστικό διεργασίας και την τελευταία διαδικασία στο αρχείο

→ χαρακτηριστικά για κάθε διαδικασία:ημερομηνία ολοκλήρωσης,οργανισμός υπαλλήλου,χρονοσφραγίδα,κατάσταση ολοκλήρωσης,αναγνωριστικό διεργασίας,υπεύθυνος ελέγχου,ονομασία στα Ολλανδικά και Αγγλικά

→ Ταξινομητές:με βάση τον κωδικό της ενέργειας(όνομα και κατάσταση),με βάση το όνομα στα Ολλανδικά(όνομα και κατάσταση),με βάση το όνομα στα Αγγλικά(όνομα και κατάσταση) και με βάση τον αρμόδιο υπάλληλο.

Summary

Από την καρτέλα Summary μπορούμε να διαπιστώσουμε για όλες τις δραστηριότητες,τον αριθμό εμφάνισης τους στο αρχείο καταγραφής συμβάντων.

1ο αρχείο

Event classes defined by Activity classifier		
All events		
Total number of classes: 398		
Class	Occurrences (absolute)	Occurrences (relative)
01_HOOFD_010+complete	1199	2,296%
01_HOOFD_015+complete	1199	2,296%
01_HOOFD_020+complete	1194	2,287%
01_HOOFD_180+complete	1116	2,137%
01_HOOFD_030_1+complete	1111	2,128%

2ο αρχείο

Event classes defined by Activity classifier		
All events		
Total number of classes: 410		
Class	Occurrences (absolute)	Occurrences (relative)
01_HOOFD_010+complete	830	1,871%
01_HOOFD_015+complete	829	1,869%
01_HOOFD_020+complete	828	1,867%
01_HOOFD_180+complete	801	1,806%
01_HOOFD_030_1+complete	792	1,786%

3ο αρχείο

Activity classifier		
Event classes defined by Activity classifier		
All events		
Total number of classes: 383		
Class	Occurrences (absolute)	Occurrences (relative)
01_HOOFD_010+complete	1409	2,361%
01_HOOFD_015+complete	1409	2,361%
01_HOOFD_020+complete	1408	2,359%

4ο αρχείο

Activity classifier		
Event classes defined by Activity classifier		
All events		
Total number of classes: 356		
Class	Occurrences (absolute)	Occurrences (relative)
01_HOOFD_010+complete	1052	2,224%
01_HOOFD_015+complete	1052	2,224%
01_HOOFD_020+complete	1051	2,222%

5ο αρχείο

Activity classifier		
Event classes defined by Activity classifier		
All events		
Total number of classes: 389		
Class	Occurrences (absolute)	Occurrences (relative)
01_HOOFD_180+complete	1161	1,965%
01_HOOFD_020+complete	1155	1,955%
01_HOOFD_015+complete	1155	1,955%
01_HOOFD_010+complete	1154	1,953%

Εικόνα 48: Καρτέλα Log Summary των 5 δήμων

Από τις παραπάνω φωτογραφίες μπορούμε να δούμε για κάθε αρχείο, ποία ήταν η πιο συνηθισμένη εργασία που εκτελέστηκε πολλές φορές, τον αριθμό εμφανίσεων της και τι ποσοστό συχνότητας είχε.

Είναι ευδιάκριτο από τα αποτελέσματα ότι όλα τα αρχεία έχουν ένα κοινό στοιχείο, κυρίως τα αρχεία 1,2,3 και 4, έχουν ως πιο δημοφιλή εργασία, την ίδια εργασία και στους 4 δήμους. Συγκεκριμένα πρόκειται για την εργασία με κωδικό 01_HOOFD_010, η οποία βρίσκεται σε ολοκληρωμένη κατάσταση, στην πρώτη θέση εμφανίσεων στους 4 δήμους. Ακόμα και στον 5ο δήμο, για λίγες μονάδες, δεν βρίσκεται στην πρώτη θέση, αλλά συναντάτε στην πρώτη τετράδα εργασιών. Αυτό μας βάζει σε σκέψεις για την εργασία αυτή και για την συμπεριφορά της, για αυτό

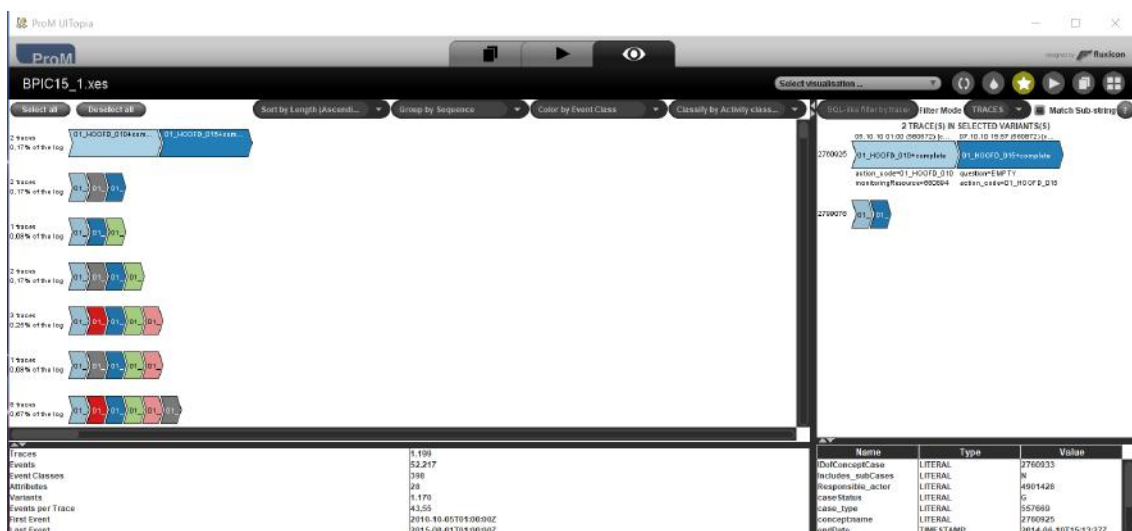
αργότερα μέσω του φιλτραρίσματος των αρχείων,θα ασχοληθούμε περισσότερο μαζί της για να δούμε πως λειτουργεί και τα μοτίβα που αναπτύσσει.Αντίστοιχα μπορούμε να αναζητήσουμε τις πιο σπάνιες εργασίες ή υπαλλήλους και να τις μελετήσουμε.Ακόμα αν θέλουμε μέσω της κατάστασης των εργασιών που αναγράφεται δίπλα από κάθε ενέργεια,έχουμε την δυνατότητα να ελέγξουμε αν κάποια διεργασία ολοκληρώνεται ή όχι.

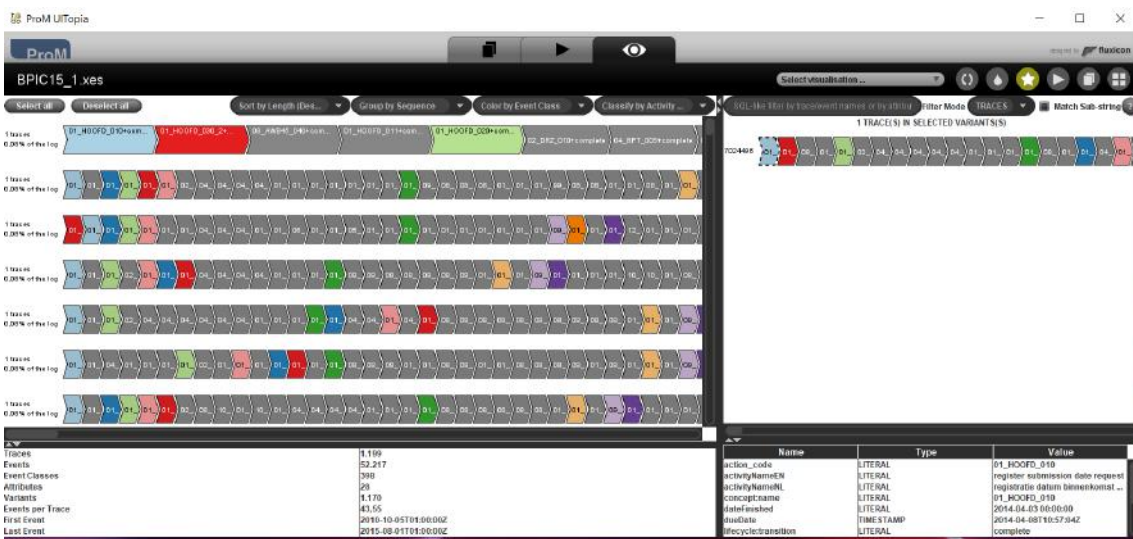
Ουσιαστικά μέσα από τις τεχνικές εικονικοποίησης που αναφερόμαστε,προσπαθούμε να παρατηρήσουμε τα δεδομένα και να βγάλουμε κάποια συμπεράσματα,τα οποία αργότερα θα φιλτράρουμε αναλυτικότερα.Έτσι και η εξόρυξη διεργασιών σαν κλάδος προσπαθεί να βρει κάποια παθολογία ή ανομοιογένεια στα δεδομένα,όπου αργότερα να την επεξεργαστεί και να κατανοήσει πως προκύπτει.Στην περίπτωση που πρόκειται για μία προβληματική συμπεριφορά,θα είναι σε θέση να την βελτιώσει,αυξάνοντας τον δείκτη αποδοτικότητας.

Explore Event Log (Trace variants/Searchable/Sortable)(Log Enhancement)

Ακόμα όπως αναφέραμε,ένα πολύ χρήσιμο εργαλείο visualization του Prom είναι το Explore event log το οποίο δείχνει όλα τα traces και τα event που περιέχονται σε καθένα από αυτά.Με μία ταξινόμηση αυτών των στοιχείων,με βάση το μήκος, μπορούμε να δούμε ποία και πόσα traces έχουν τα λιγότερα event και ποία τα περισσότερα.Είναι ένα μέτρο που μπορεί να μας δείξει σε ποιές περιπτώσεις έχουμε τους πιο αποδοτικούς εργαζομένους και σε ποιές τους λιγότερα αποδοτικούς.

1ο αρχείο

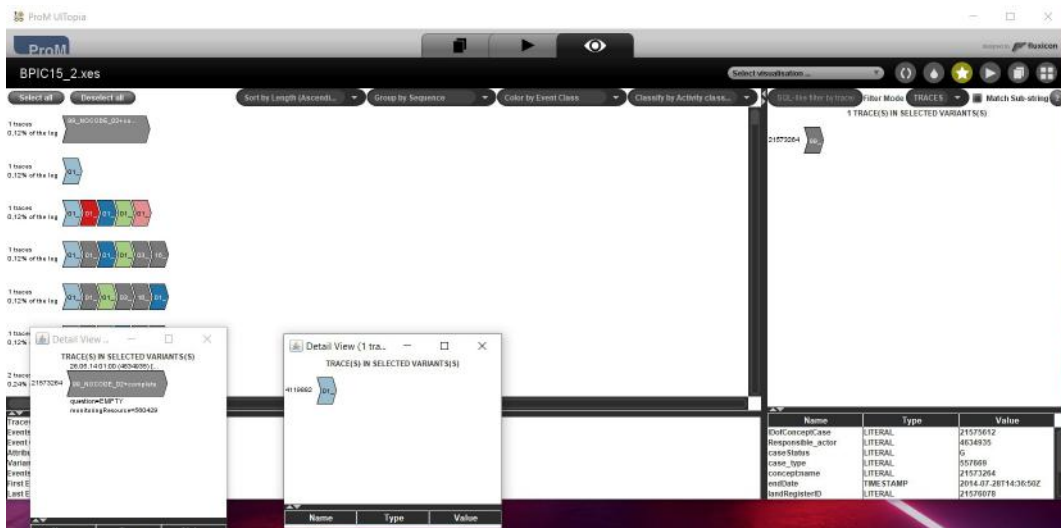




Εικόνα 49: Explore Event Log 1ου δήμου

Βλέπουμε ότι έχουμε 2 υποθέσεις με τον μικρότερο αριθμό από δραστηριότητες και για την ακρίβεια 2 event ο καθένας. Στη συνέχεια παρατηρούμε ότι τα επόμενα traces κλιμακώνονται σταδιακά, αυξάνοντας τον αριθμό από event κατά μία μονάδα.

2ο αρχείο



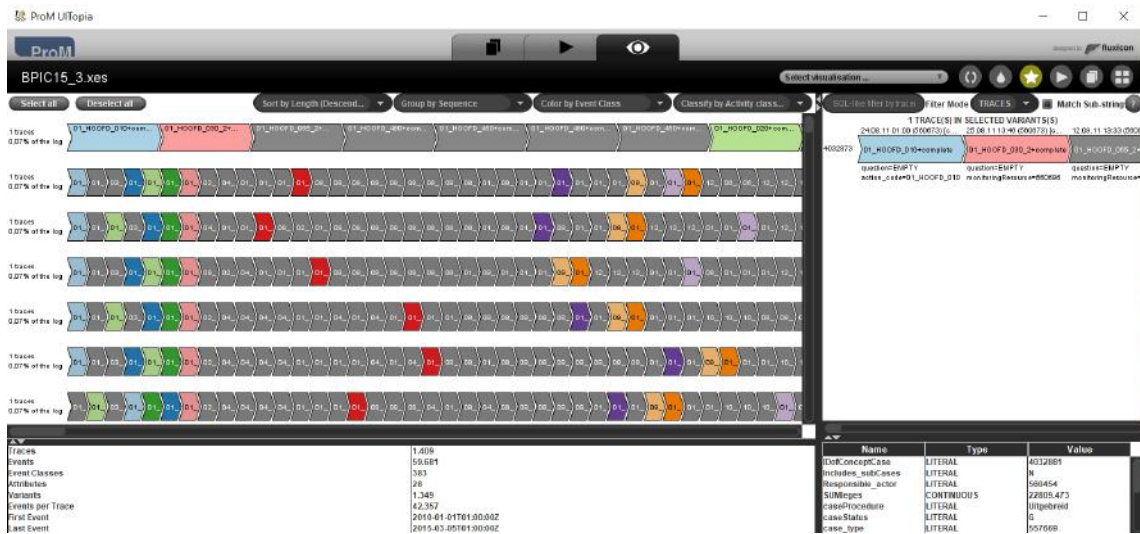


Εικόνα 50: Explore Event Log 2ου δήμου

Εδώ αντιλαμβανόμαστε ότι έχουμε 2 traces με μόλις 1 event υλοποιημένο που τα καθιστά από τα λιγότερα αποδοτικά traces και θα έπρεπε να εστιάζει εκεί ο 2ος δήμος. Παρόλα αυτά τα αμέσως επόμενα traces περιλαμβάνουν από 5 και πάνω events, κάνοντας μας να σκεφτούμε το πρόβλημα έγκειται στα 2 πρώτα traces ενώ μετά εξομαλύνεται η κατάσταση.

3ο αρχείο

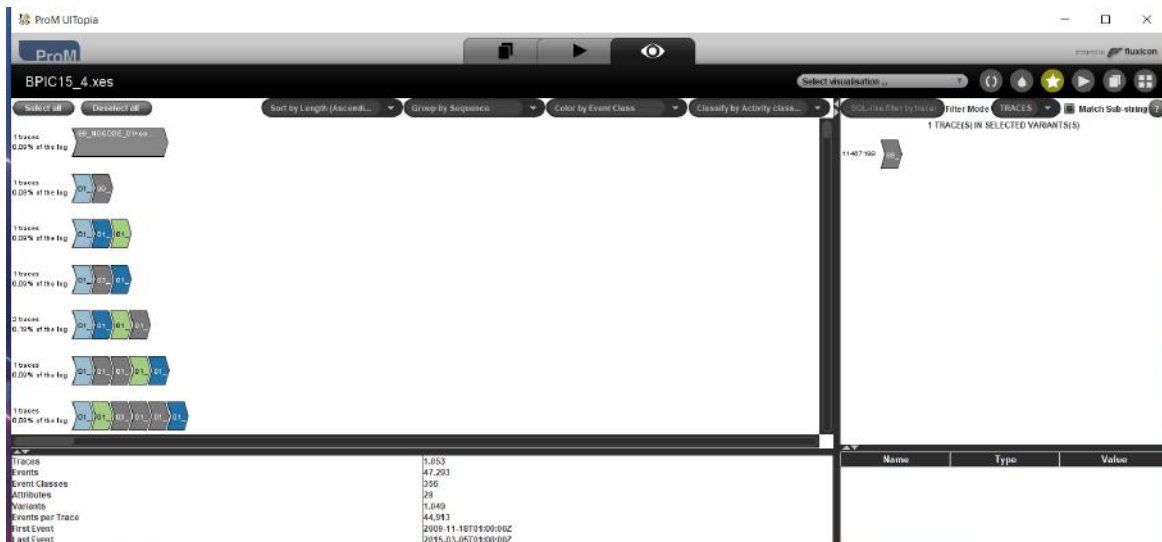




Εικόνα 51: Explore Event Log Ζου δήμου

Ο Ζος δήμος βλέπουμε να ξεκινάει με καλύτερα αποτελέσματα καθώς υπάρχουν 4 trace με 3 event και όλα τα επόμενα trace περιλαμβάνουν από 4 και πάνω event.

4ο αρχείο

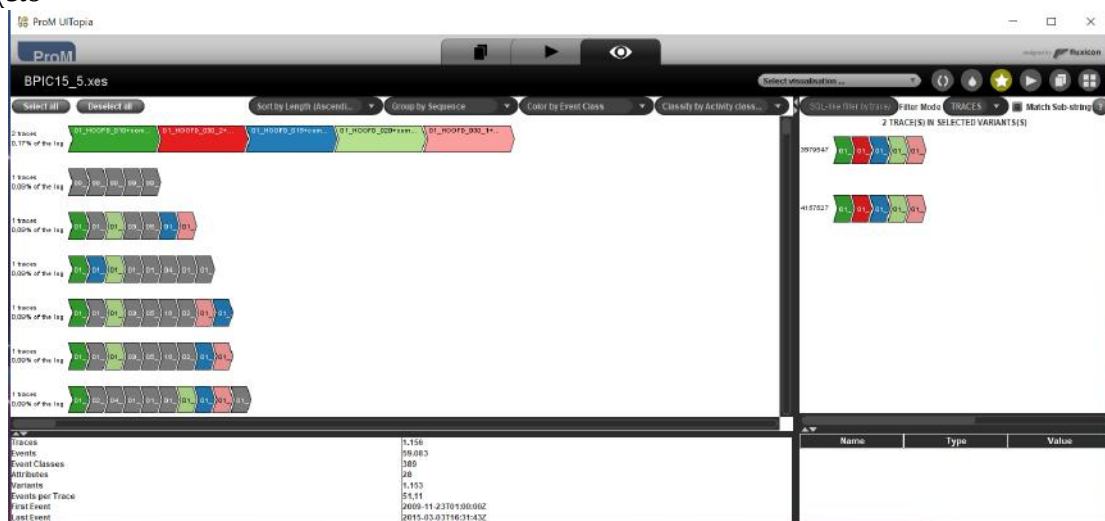


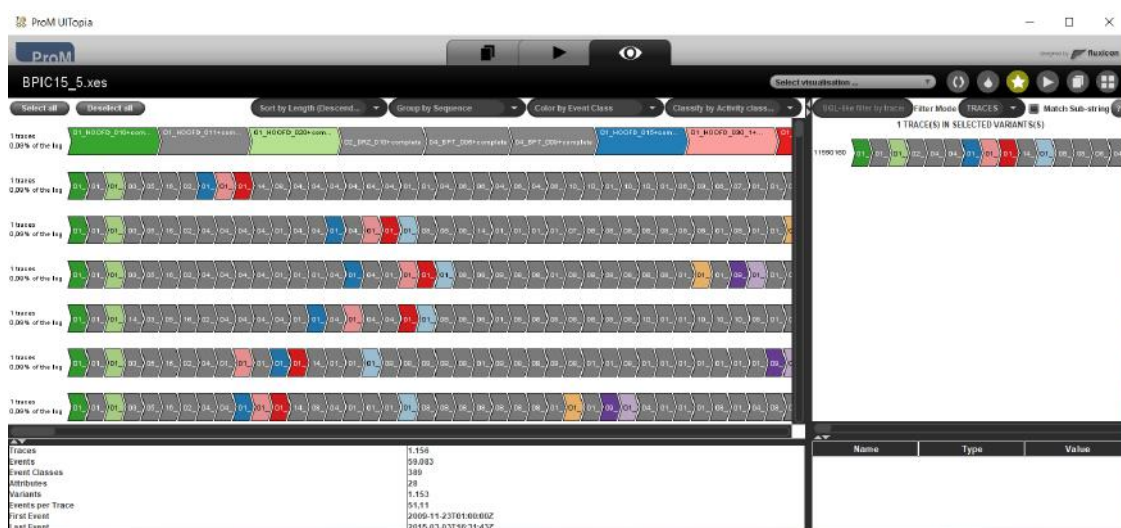


Εικόνα 52: Explore Event Log 4ου δήμου

Στο 4ο δήμο συναντάμε μία αντίστοιχη περίπτωση με τον 2ο δήμο όπου βλέπουμε trace με μόλις 1 event και μετά παρατηρούμε μία κλιμακωτή αύξηση των δραστηριοτήτων. Όμως χωρίς να πιάνει τις επιδόσεις του 2ου δήμου όπου μετά τα πρώτα ανησυχητικά trace, τα επόμενα είχαν πολλές δραστηριότητες, 5 και πάνω συγκεκριμένα.

5ο αρχείο





Εικόνα 53: Explore Event Log 5ου δήμου

Στο 5ο και τελευταίο αρχείο συναντάμε την ίσως καλύτερη απόδοση από τους 5 δήμους. Συγκεκριμένα τα trace με τα λιγότερα event ξεκινάνε από τα 5 events και αυξάνονται αρκετά, σε αντίθεση με τους υπόλοιπους δήμους όπου αρκετά trace είχαν πολύ λιγότερα από 5 event έκαστως. Αναλογιζόμενοι και την πληθώρα των event που εκτελούνται στον δήμο, με σχεδόν 60.000 event, παρατηρούμε ότι έχει την καλύτερη απόδοση και παραγωγή μέχρι στιγμής.

Στη συνέχεια παρατηρώντας το ίδιο plug in που είχαμε και παραπάνω μπορούμε να δούμε για κάθε αρχείο ποιος είναι ο μέσος όρος δραστηριοτήτων ανά υπόθεση. Αυτό βοηθάει την έρευνα μας για κάθε αρχείο των δήμων βλέποντας, με τον μέσο όρο πάντα, ποιος δήμος εκτελούσε τις περισσότερες εργασίες.

Παρατηρούμε λοιπόν για το: 1ο αρχείο 43,55 events per trace

→ 53,31 για το 2ο αρχείο

→ 42,357 για το 3ο αρχείο

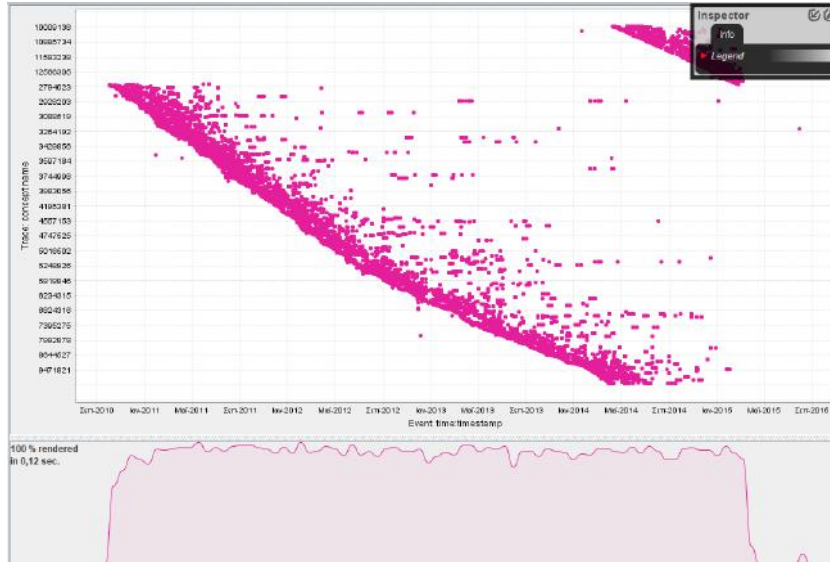
→ 44,913 για το 4ο αρχείο

→ 51,11 για το 5ο αρχείο

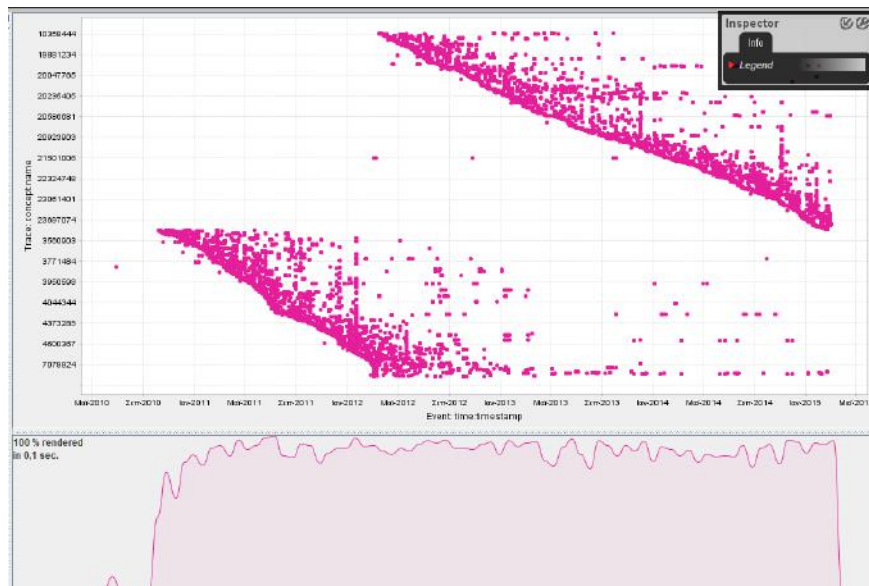
Βλέπουμε σε αυτήν την περίπτωση τον αρκετά υψηλό μέσο όρο του 2ου και 5ου δήμου, καθώς ξεπερνάνε τα 50 events στο μέσο όρο. Επισημαίνοντας ακόμα ότι ο 5ος δήμος μας έχει φέρει σε πολλές μετρήσεις αξιόλογα αποτελέσματα. Στη συνέχεια οι δήμοι 1,3,4 έχουν χαμηλότερη απόδοση, καθώς κυμαίνονται περίπου στα 43 events μέσο όρο ανά trace. Όχι και τόσο καλή απόδοση αν αναλογιστούμε ότι ο 1ος και ο 3ος δήμος ξεπερνούν τα 50.000 events. Ενώ από την άλλη πλευρά ο 2ος δήμος με μόλις 44.354 events συγκέντρωσε τον υψηλότερο μέσο όρο. [20]

Dotted Chart

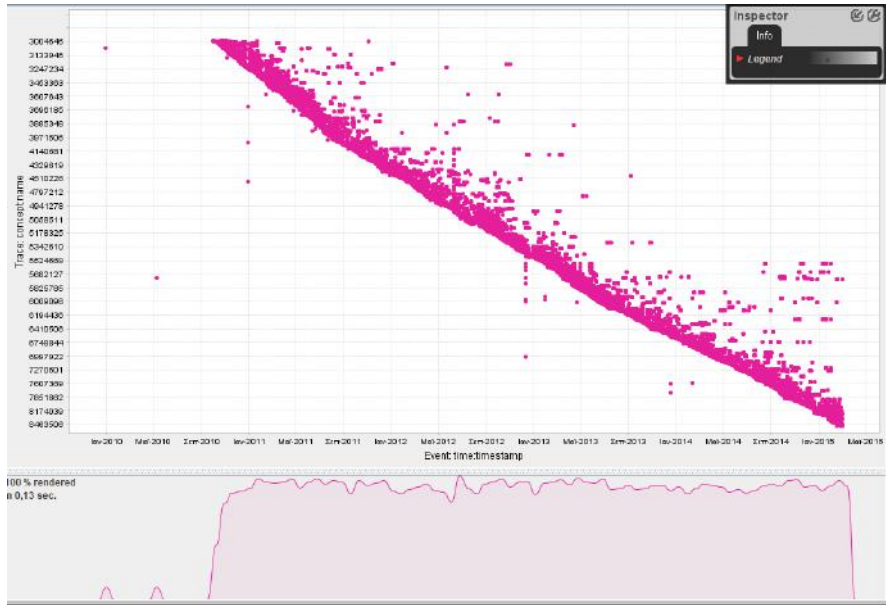
Επιπλέον μία χρήσιμη τεχνική εικονικοποίησης που προσφέρει το Prom Tools είναι το dotted chart. Η τεχνική αυτή δείχνει την χρονική σήμανση των γεγονότων και σε ποιές περιόδους είχαμε υψηλή ή χαμηλή απόδοση στις εργασίες. Όσο πιο πυκνό το διάγραμμα τόσο μεγαλύτερη η απόδοση. 1ο αρχείο



2ο αρχείο



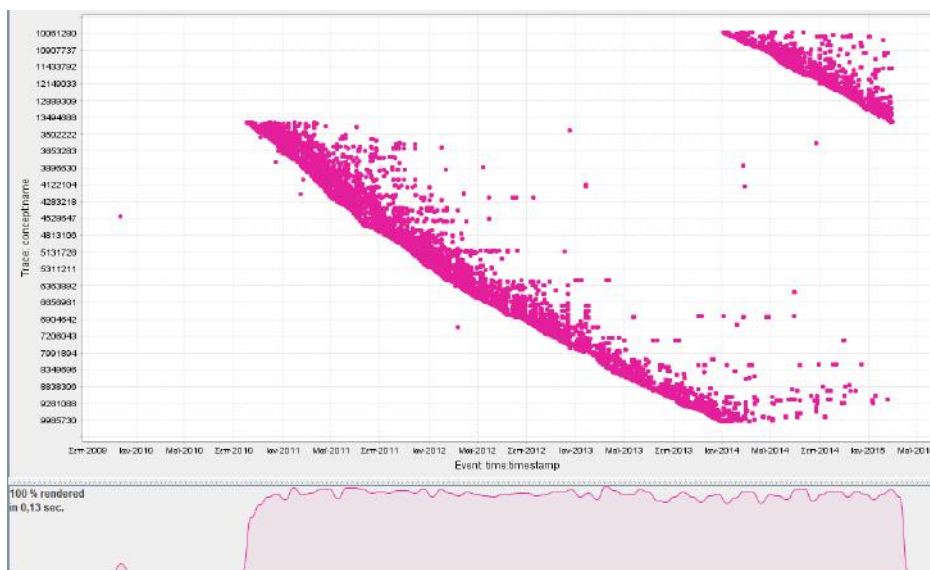
3ο αρχείο



4ο αρχείο



5ο αρχείο



Εικόνα 54: Dotted Chart των δήμων

Τα παραπάνω dotted chart διαγράμματα μας δείχνουν στον κάθετο άξονα τα cases και στον οριζόντιο άξονα τον χρόνο,δειγνοντας τα activities που εκτελεί ο κάθε υπάλληλος σε συνάρτηση με το χρόνο.Οι μετρήσεις που βλέπουμε ξεκινάνε για όλα τα αρχεία από τα τέλη του 2009 και στη συνέχεια μέσα στο 2010 πληθαίνουν.Εξάιρεση αποτελεί ο 1ος δήμος που περιέχει αρκετές εργασίες ήδη από το 2009.Βλέπουμε στα αρχεία 1,2 και 5 πυκνότερα διαγράμματα ,που περιέχουν πολλές περισσότερες εργασίες σε σχέση με τους υπόλοιπους δήμους.Ακόμα οι κυματομορφές κάτω από τα διαγράμματα δείχνουν τα πολυπληθή cases των δήμων 1,2 και 5,οι οποίες παραμένουν σε υψηλό δείκτη όλη την διάρκεια της 4ετίας,σε αντίθεση με τους υπόλοιπους δήμους που φτάνουν σε ένα peak της απόδοσης τους και μετά υποχωρούν..Διατηρούν μία πιο ομοιόμορφη απόδοση σε βάθος της 4ετίας,διατηρώντας την σε αρκετά υψηλά επίπεδα.Αξιοσημείωτο σημείο φαίνεται να είναι ότι σε όλα τα αρχεία η κορύφωση των εργασιών σημειώνεται περίπου το πρώτο 6μηνο του 2013,όπου εκεί φαίνεται να ήταν η πιο παραγωγική περίοδος για τους δήμους.Ακόμα πρέπει να επισημάνουμε ότι σε όλους τους δήμους έχουμε υποθέσεις,οι οποίες ξεκινούν τις διαδικασίες τους αργότερα από το αναμενόμενο,μερικές μετά την μέση της 4ετίας.Μέσα από τα παραπάνω διαγράμματα μπορούμε ξανά να παρατηρήσουμε την συμπεριφορά των ενεργειών και να εξάγουμε κάποια συμπεράσματα για αυτές,ούτως ώστε στην συνέχεια να φιλτράρουμε αυτές τις διαδικασίες και να τις μελετήσουμε.

4.1.3 PROCESS MODEL

Όπως προαναφέραμε στο Prom Tools έχουμε αρκετά plug in διαθέσιμα για την δημιουργία των μοντέλων διαδικασιών.Με αυτήν την τεχνική εφαρμόζουμε έναν αλγόριθμο που επιθυμούμε στο αρχείο και βλέπουμε το παραχθέν μοντέλο διαδικασιών.Έτσι είμαστε σε θέση να αντιληφθούμε κατά πόσο συμφωνούν τα στοιχεία που έχουμε μέχρι τώρα από τις μετρήσεις ,με την αναπαράσταση του διαγράμματος των εργασιών.Εφαρμόζοντας αυτούς τους αλγόριθμους περνάμε

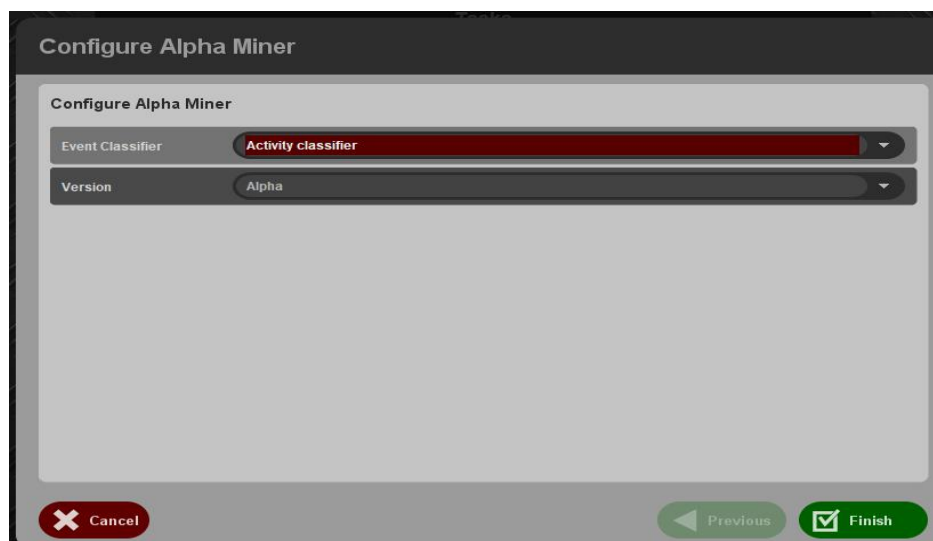
ουσιαστικά από την θεωρία που είδαμε μέχρι τώρα,στην πράξη όπου μπορούμε να παρατηρήσουμε τον τρόπο που λειτουργούν οι διαδικασίες και πως αλληλεπιδρούν μεταξύ τους.

Πολλές φορές στις δοκιμές μας,ιδιαίτερα όταν έχουμε να κάνουμε με ένα αξιολογούμενο όγκο δεδομένων επιδιώκουμε να κάνουμε πρώτα φιλτράρισμα ενός αρχείου,με βάση κάποιο κριτήριο και μετά να εφαρμόσουμε στο παραγόμενο αρχείο έναν αλγόριθμο.Συνήθως στην αρχική είσοδο ,το μοντέλο διαδικασιών που παράγεται είναι δυσνόητο,γιατί περιέχει αρκετή πληροφορία.Για αυτούς του λόγους συνήθως ,πρώτα φιλτράρουμε ένα αρχείο και έπειτα χρησιμοποιούμε τον επιθυμητό αλγόριθμο.Έτσι βλέπουμε αν τα στοιχεία που πήραμε από το φιλτράρισμα,είναι εμφανή και στο μοντέλο μας.

Τα plug in που έχουμε στη διάθεση μας για τους αλγόριθμους mining είναι τα εξής:

- **Alpha miner** για τον alpha αλγόριθμο
- **Mine fr a heuristics net using heuristics miner** για τον heuristic αλγόριθμο
- **Mine with inductive visual miner** για τον inductive αλγόριθμο
- **Mine for a fuzzy model** για τον fuzzy αλγόριθμο

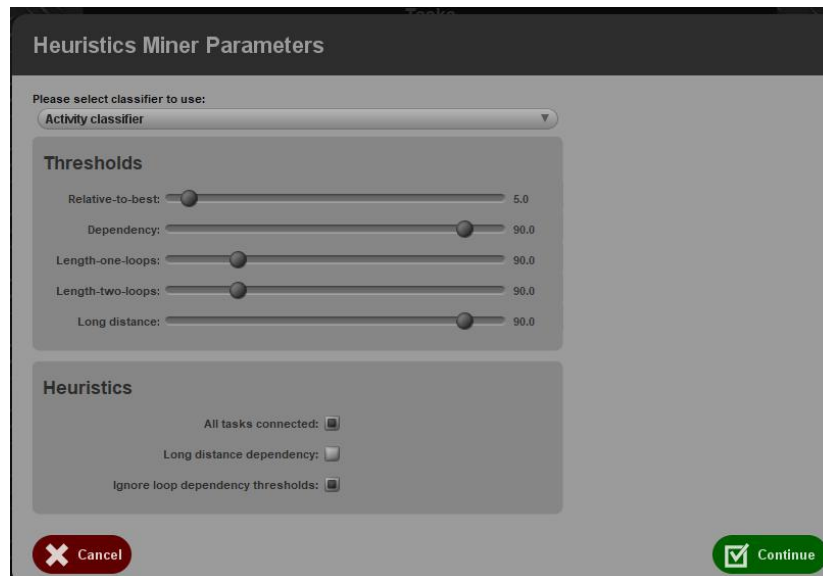
Alpha



Εικόνα 55: Alpha miner plug in

Ο Alpha miner δεν είναι ένας πολύ καλός αλγόριθμος για την εξόρυξη διεργασιών, αλλά μερικοί αποτελεί την ρίζα των μετέπειτα σύγχρονων αλγορίθμων,οπότε θα τον χρησιμοποιήσουμε.Μπορούμε να επιλέξουμε τον διαχωριστή με βάση τον οποίο θα εμφανιστούν τα αποτελέσματα,στο παράδειγμα μας έχουμε 4 διαχωριστές και επιλέγουμε αυτόν που μας δίνει την δραστηριότητα και ως έκδοση την Alpha.Το Prom διαθέτει και άλλες παραλλαγές όπως (alpha+, alpha++, alpha#, alphaR) ανάλογα τι επιθυμούμε και το προκύπτον μοντέλο διαδικασίας παρουσιάζεται ως petri net.

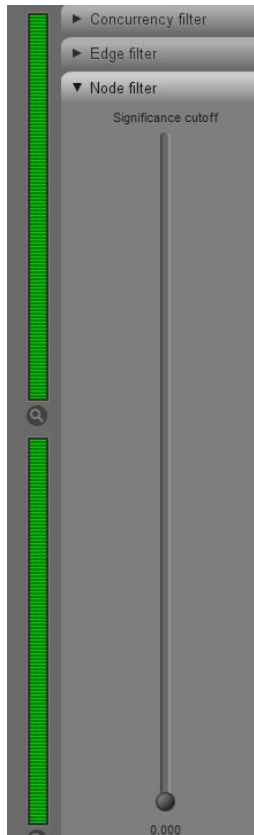
Heuristic



Εικόνα 56: Heuristic miner plug in

Όπως αναλύσαμε προηγουμένως ο heuristic miner λαμβάνει υπόψη τις συχνότητες και έτσι μπορεί να φιλτράρει τη σπάνια συμπεριφορά. Επίσης, είναι σε θέση να ανιχνεύσει βρόχους και παραλείψεις. Έχουμε διάφορες παραμέτρους που μπορούμε να ελέγξουμε και κυρίως την παράμετρο "Thresholds" "Σχετικά με το καλύτερο" που υποδηλώνει ότι μόνο συμβάντα που παρατηρούνται περισσότερες φορές από την ορισμένη τιμή κατωφλίου θα ληφθούν υπόψη. Συγκεκριμένα μόνο γεγονότα που έχουν παρατηρηθεί πάνω από 5 φορές θα συμπεριληφθούν στο παραγόμενο μοντέλο. Ακόμα ελέγχουμε τις τιμές κατωφλίου για το μήκος των βρόχων και των αποστάσεων μεταξύ των διεργασιών και αν τα γεγονότα θα είναι συνδεδεμένα ή θα παραλειφθούν από το μοντέλο. Τέλος όπως και προηγουμένως μπορούμε να επιλέξουμε τον επιθυμητό διαχωριστή, ανάλογα με την δοκιμή που κάνουμε.

Fuzzy miner



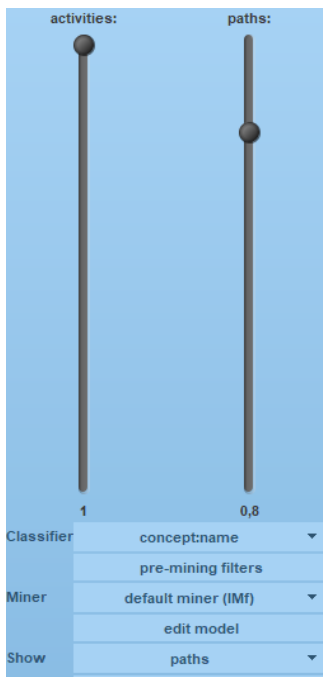
Στην περίπτωση του Fuzzy miner διαθέτουμε 3 ειδών φίλτρα, το node filter ρυθμίζει το πλήθος των διαφορετικών διαδικασιών που θα παρουσιαστούν στο γράφημα. Μπορούμε να ελέγξουμε το ποσοστό αποκοπής και έτσι κάθε διαδικασία με ποσοστό εμφάνισης χαμηλότερο από το κατώφλι που θέτουμε, θα αφαιρεθεί από το μοντέλο. Σε άλλη περίπτωση θα συμπεριληφθεί στο παραγόμενο γράφημα.

Το edge filter επηρεάζει την ποσότητα των ακμών, και την επιλογή τους, η οποία θα περιλαμβάνεται στο τελικό γράφημα. Η δοθείσα τιμή επηρεάζει το ποσοστό των ακμών που θα περιληφθούν στο διάγραμμα. Όσο μεγαλύτερη είναι αυτή η τιμή τόσο περισσότερες ακμές θα συμπεριλαμβάνονται. Ενώ αντίστοιχα όσο μικρότερη η τιμή τόσο λιγότερες ακμές στο διάγραμμα, διατηρώντας μία αναλογική σχέση.

Τέλος το concurrency filter έχει την χρησιμότητα του στον συγχρονισμό που θέλουμε να εφαρμόσουμε. Μπορεί να φιλτράρει τις συγκρούσεις που προκύπτουν μεταξύ δύο ή περισσότερων δραστηριοτήτων. Με τον όρο σύγκρουση αναφερόμαστε στην κατάσταση όπου 2 κόμβοι (ή δραστηριότητες) συνδέονται αμφίδρομα μεταξύ τους, το οποίο μπορεί να προκαλέσει βρόχους και αδιέξοδα στο μοντέλο μας.

Εικόνα 57: Fuzzy miner plug in

Inductive visual miner



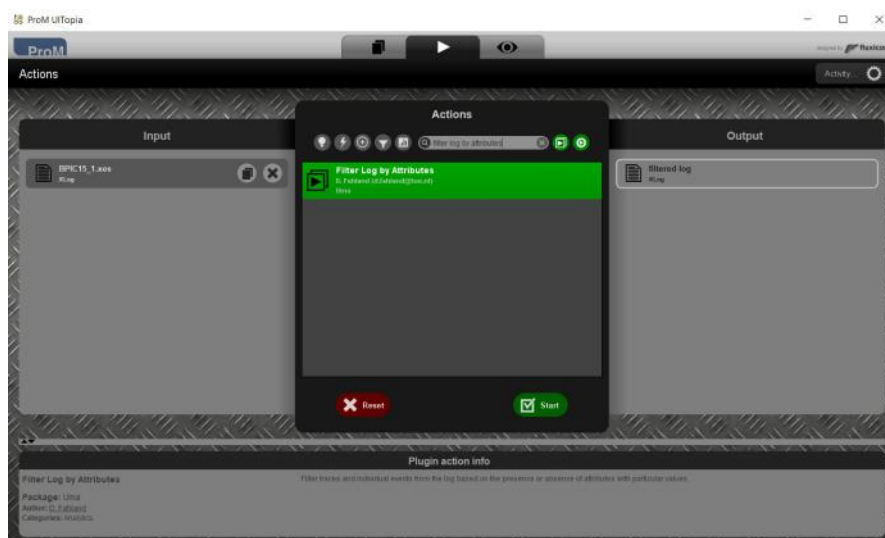
Στον inductive visual miner μπορούμε να παραμετροποιήσουμε διάφορες τιμές, όπως τη συσχέτιση μεταξύ των τιμών για τις δραστηριότητες και τα μονοπάτια που θα εμφανίζονται. Αυξάνοντας αυτές τις τιμές τοποθετούμε όλο και περισσότερα activities στο διάγραμμα, ενώ με την μείωση τους κάνουμε το μοντέλο πιο απλοϊκό. Ακόμα αποφασίζουμε και εδώ τον διαχωριστή που θα χρησιμοποιήσουμε και ποίο στοιχείο (attribute) θα εμφανίζεται. Επίσης μπορούμε να προσθέσουμε και κάποια επιπλέον φίλτρα για την εξόρυξη του μοντέλου και την επεξεργασία του, μετά από τον χρήστη. Τέλος ελέγχουμε αν θα εμφανίζει και επιπλέον πληροφορίες, εκτός από τις δραστηριότητες και τα μονοπάτια. [21]

Εικόνα 58: Inductive Visual miner plug in

4.1.4 FILTERING EVENT LOG

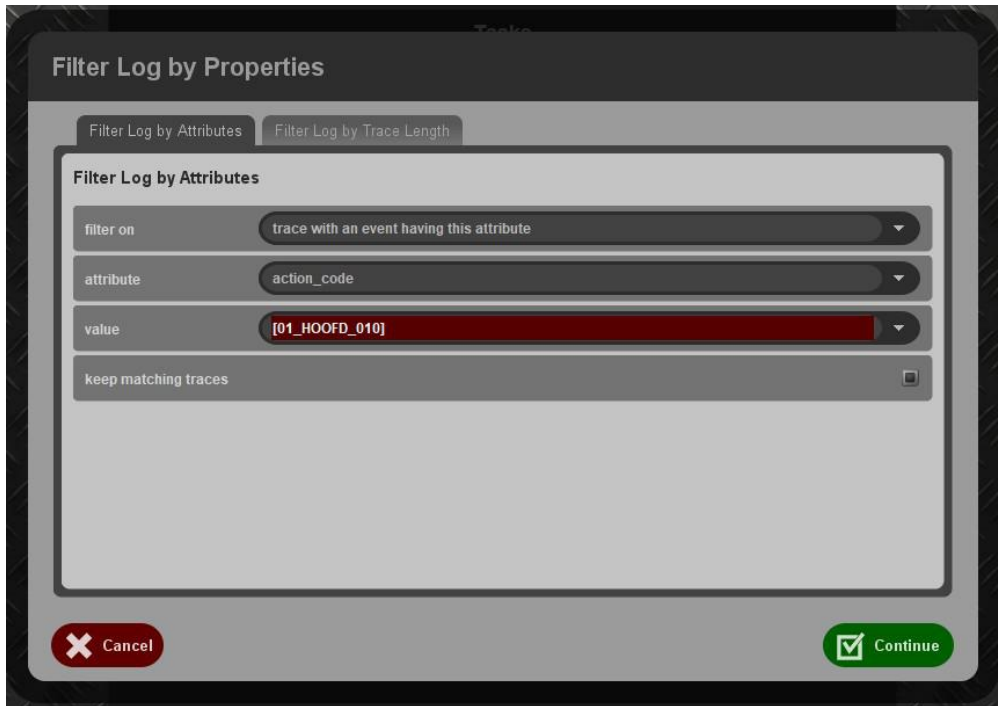
Μέσα από την πλατφόρμα του Prom Tools έχουμε πρόσβαση σε αρκετά διαθέσιμα plug in για φιλτράρισμα των αρχείων καταγραφής συμβάντων. Παρατηρώντας κάποια ανομοιογένεια ή ευπάθεια από προηγούμενες μετρήσεις, αυτήν την στιγμή μπορούμε να την μελετήσουμε διεξοδικά, μέσω των διαθέσιμων εργαλείων. Όπως προαναφέραμε έχουμε την δυνατότητα να ασχοληθούμε με εργασίες με υψηλή συχνότητα εμφάνισης ή αντίστοιχα με εργασίες που εκτελούνται αρκετά σπάνια. Μπορούμε να εφαρμόσουμε τεχνικές εξόρυξης στους εργαζομένους των δήμων για να ανακαλύψουμε σε ποιές περιπτώσεις έχουμε τους πιο πολυάσχολους υπαλλήλους. Ακόμα θα μελετήσουμε την συμπεριφορά των δήμων όσον αφορά τους εργαζομένους για να βρούμε τους δήμους με την υψηλότερη παραγωγικότητα. Σε αρκετές δοκιμές μετά την εφαρμογή του φίλτρου και την παραγωγή του νέου αρχείου, θα εφαρμόσουμε αλγόριθμους δημιουργίας μοντέλου διαδικασιών, για να δούμε την νέα ροή εργασιών. Αυτούς τους αλγόριθμους μπορούμε να τους εφαρμόσουμε και στο αρχικό αρχείο εισόδου, για να ρίξουμε μία πρώτη ματιά στο αρχείο και στις διεργασίες που εκτελούνται.

Όπως αναφέραμε το Prom Tools διαθέτει αρκετά plug in για το φιλτράρισμα του αρχείου καταγραφής συμβάντων. Στις επόμενες μετρήσεις εφαρμόζουμε το plug in Filter log by attributes, με βάση την πιο συνηθισμένη δραστηριότητα με κωδικό 01_HOOFD_010 στα 5 αρχεία των δήμων. Γίνεται απεικόνιση των στοιχείων με το LogDialog/Dashboard και έπειτα χρήση alpha, heuristic, Inductive visual και fuzzy miner για αναπαράσταση του process model.



Εικόνα 59: Καρτέλα Actions για επιλογή plug in

Επιλέγουμε φιλτράρισμα στο ίχνος που περιλαμβάνει αυτήν την δραστηριότητα, με βάση τον κωδικό της ενέργειας και επιλέγουμε το επιθυμητό κωδικό από την διαθέσιμη λίστα και πατάμε Continue, για την παρουσίαση των αποτελεσμάτων.

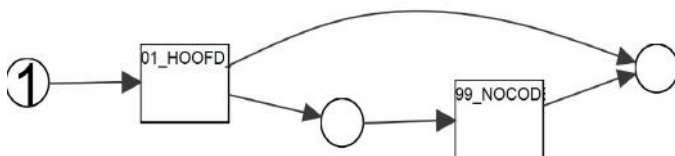


Εικόνα 60: Filter Log By attributes plug in

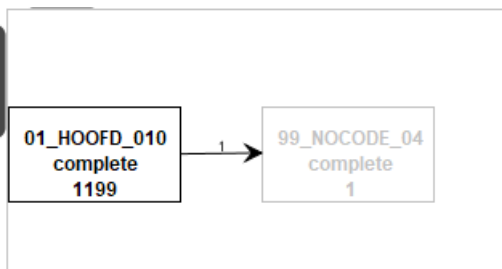
1ο αρχείο
Dashboard



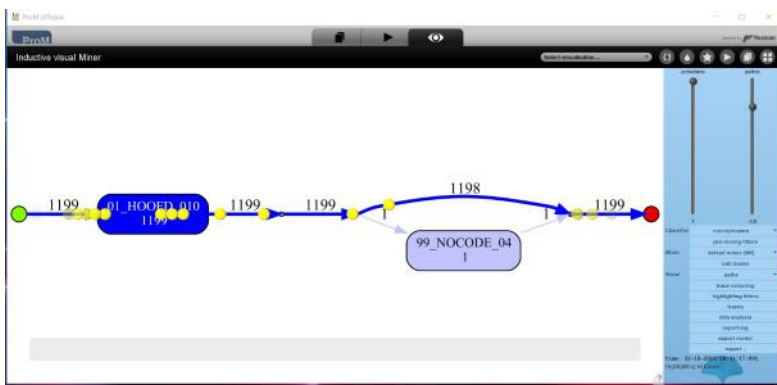
Alpha



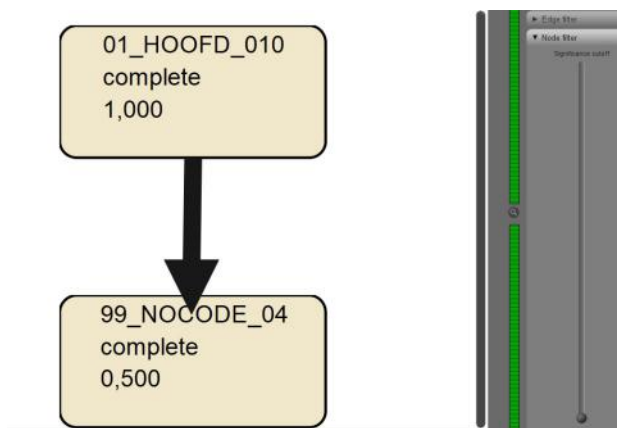
Heuristic



Inductive



Fuzzy

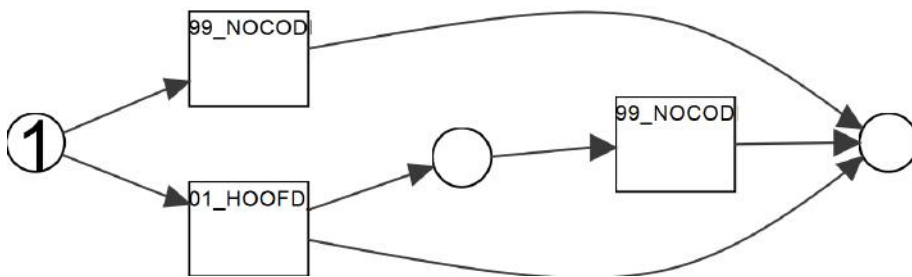


Εικόνα 61: Dashboard και Process model 1ου αρχείου για εργασία 01_HOOFD_100

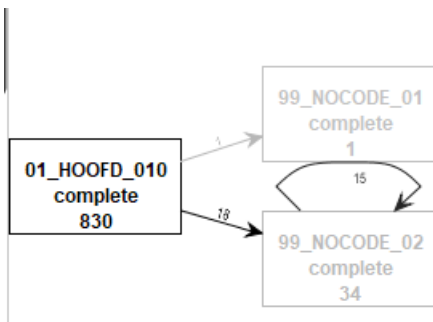
2ο αρχείο
Dashboard



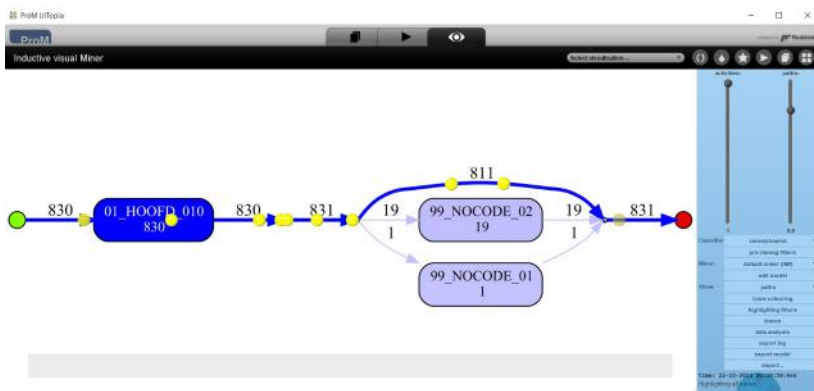
Alpha



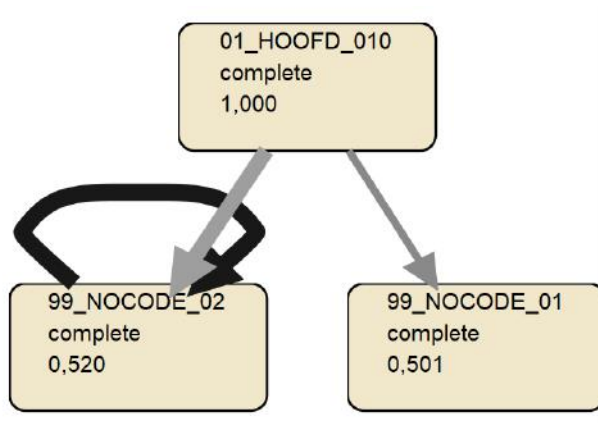
Heuristic



Inductive



Fuzzy

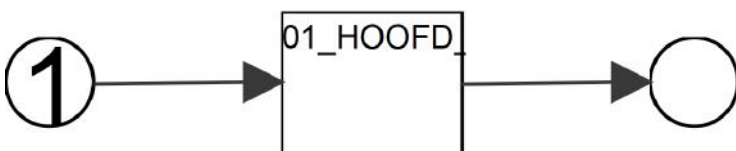


Εικόνα 62: Dashboard και Process model 2ου αρχείου για εργασία 01_HOOFD_100

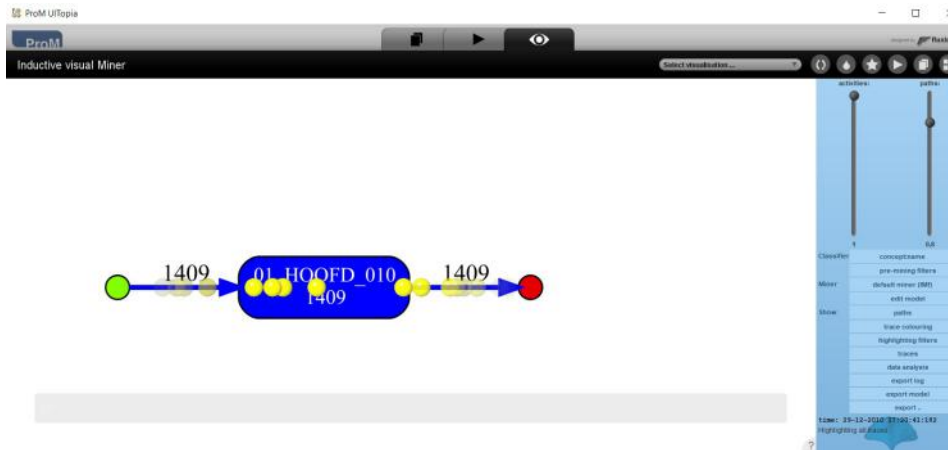
3ο αρχείο
Dashboard



Alpha

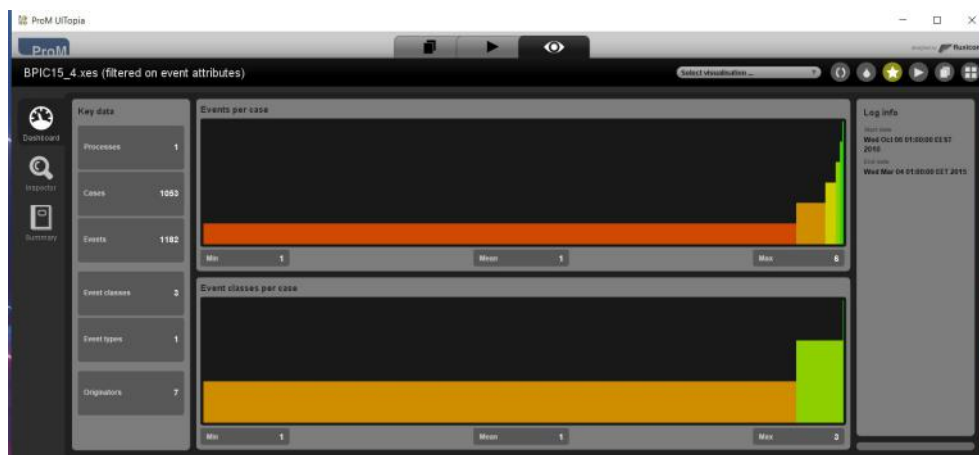


Inductive

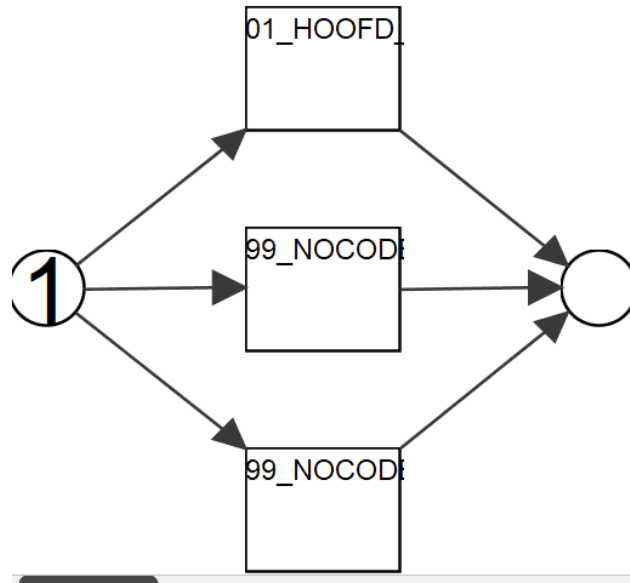


Εικόνα 63: Dashboard και Process model 3ου αρχείου για εργασία 01_HOOFD_100

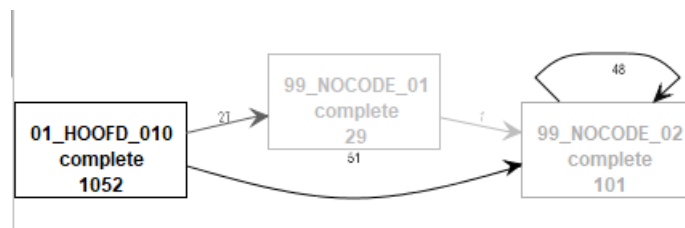
4ο αρχείο
Dashboard



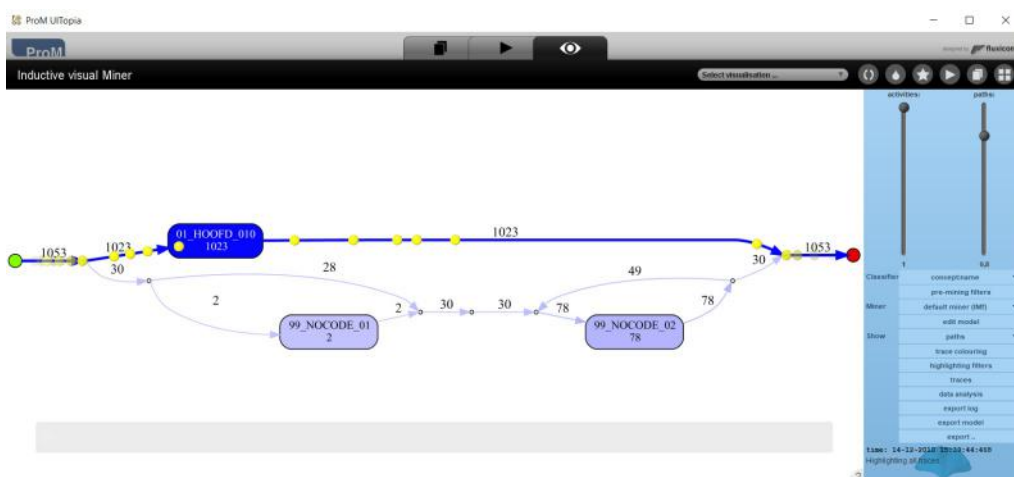
Alpha



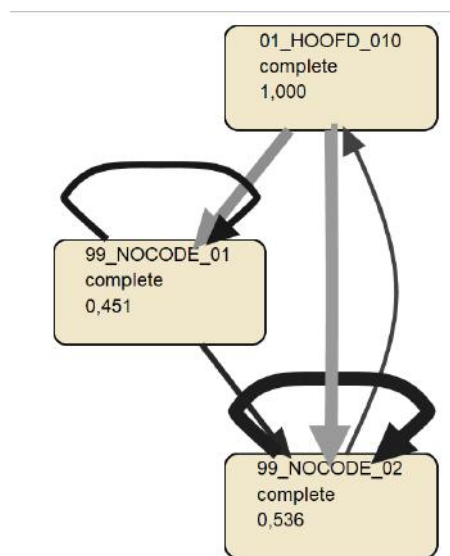
Heuristic



Inductive



Fuzzy



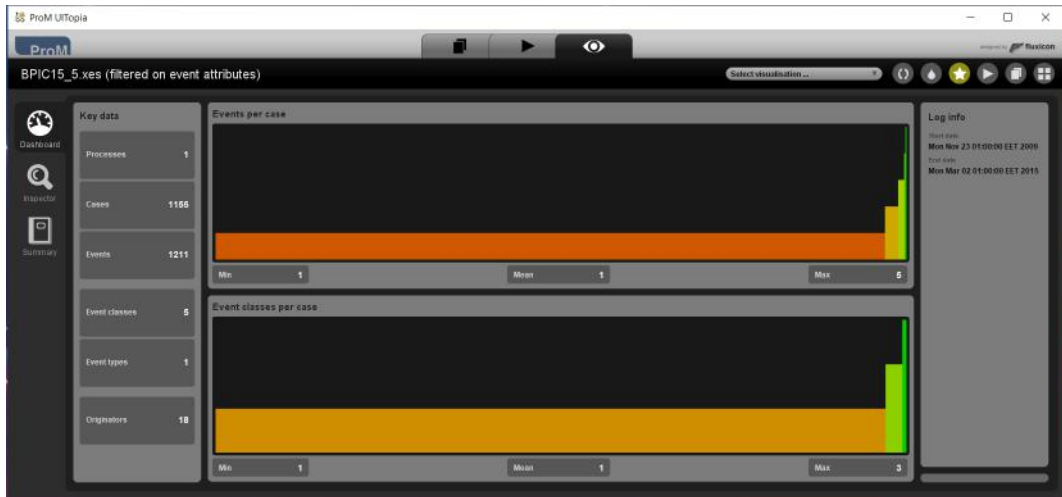
Εικόνα 64: Dashboard και Process model 4ου αρχείου για εργασία 01_HOOFD_100

Παρατηρούμε ότι στα 4 αρχεία το activity 01_HOOFD_010 εμφανίζεται με έναν μέσο όρο 1.122 φορές γεγονός που το κάνει από τα πιο δημοφιλή activity στους δήμους. Στο αρχείο νούμερο 3 έχει τον μεγαλύτερο αριθμό εμφανίσεων 1409 για την ακρίβεια, ενώ από την άλλη μεριά στο αρχείο 2 έχει τις λιγότερες εμφανίσεις με αριθμό 830.

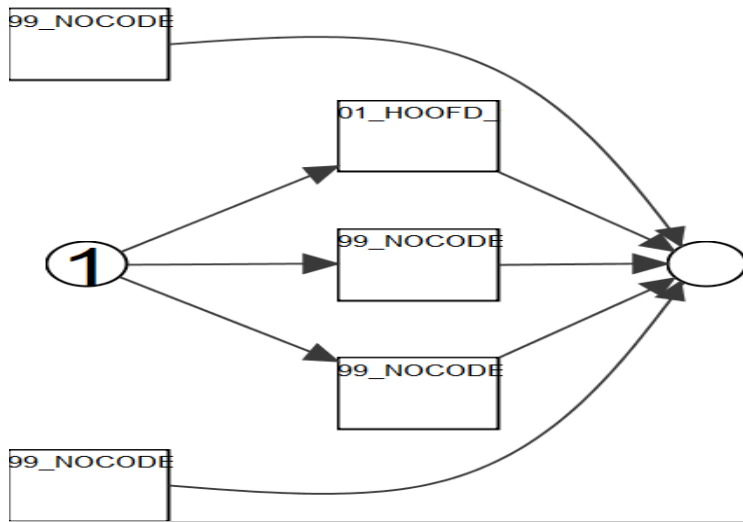
Ακόμα σημειώνουμε ότι στο 3ο αρχείο παράγεται ,μέσω των alpha,fuzzy και inductive visual miner, το απλούστερο και πιο κατανοητό process model που έχουμε δει μέχρι τώρα. Επί προσθέτως στο 3ο αρχείο έχουμε την μεγαλύτερη ομοιογένεια στον αριθμό εμφανίσεων του event σε κάθε case, για την ακρίβεια ένα τη φορά. Αντίθετα όμως τα αρχεία 2 και 4 παρουσιάζουν τα πιο πολύπλοκα και δυσνόητα process model και την μεγαλύτερη ανομοιογένεια στην κατανομή της δραστηριότητας στις υποθέσεις. Πράγμα το οποίο τα καθιστά όχι και τόσο αποδοτικά και εύρυθμα. Παρατηρούμε ότι στα παραγόμενα process models ο αλγόριθμος inductive visual miner εμφανίζει τα πιο κατανοητά και παραστατικά αποτελέσματα, πράγμα που επιβεβαιώνει την χρησιμότητα του.

Στη συνέχεια ακολουθούν και για τον 5ο δήμο τα process models , για την πιο δημοφιλή δραστηριότητα με κωδικό 01_HOOFD_180, με την χρήση των ίδιων miners όπως προηγουμένως. Και σε αυτήν την περίπτωση όμως η δραστηριότητα 01_HOOFD_010 ήταν αρκετά συχνή στις εκτελεσμένες εργασίες.

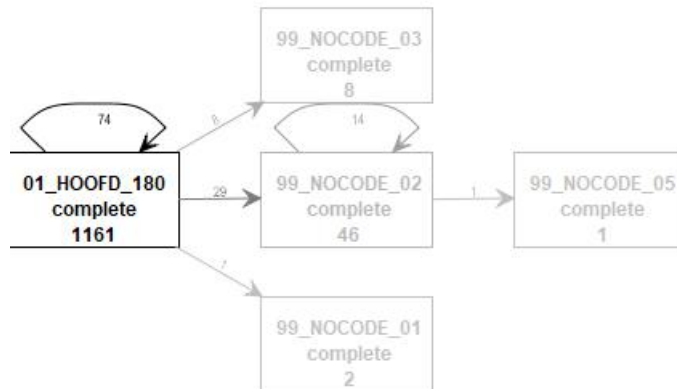
5ο αρχείο
Dashboard



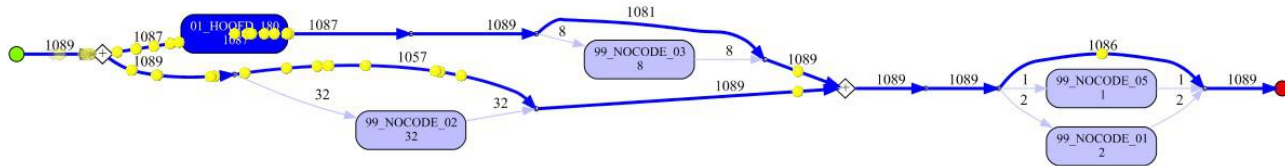
Alpha



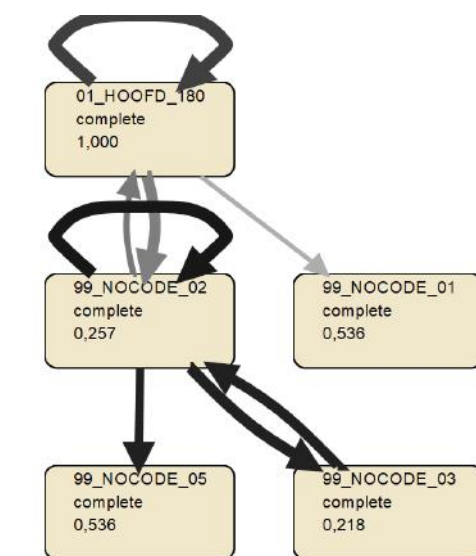
Heuristic



Inductive



Fuzzy



Εικόνα 65: Dashboard και Process model 5ου αρχείου για εργασία 01_HOOFD_180

Βλέπουμε ότι στο 5ο δήμο τα process model που παράγονται για το αντίστοιχο activity είναι πιο πολύπλοκα σε σχέση με τους υπόλοιπους δήμους και περιέχουν αρκετά μεγαλύτερη πληροφορία. Παρατηρούμε ότι έχουμε διαγράμματα με βρόχους, αμφίδρομες σχέσεις και μονοπάτια που θα μπορούσαν να παραλειφθούν. [22]

Σε συνέχεια από τις προηγούμενες μετρήσεις επιδίωξα με την χρήση του plug in filter log by attributes και συγκεκριμένα στην καρτέλα filter log by trace length, να βρω για κάθε δήμο σημεία με ιδιαίτερο ενδιαφέρον. Για αρχή έθεσα ως κατώφλι στα ίχνη, αυτά που έχουν μήκος δραστηριοτήτων κάτω από 5 και μετά στο παραγόμενο αρχείο εφαρμόσα τον αλγόριθμο Inductive visual, που όπως είδαμε αποδίδει αρκετά καλά, για να δούμε το γραφικό μοντέλο. Βασικός σκοπός ήταν να βρω σε κάθε δήμο, πόσους εργαζομένους έχουμε που αναλαμβάνουν και εκτελούν πολλοί λίγες δραστηριότητες. Με αυτόν τον τρόπο θα δούμε πόσοι υπάλληλοι εκτελούν κάτω από μόλις 5

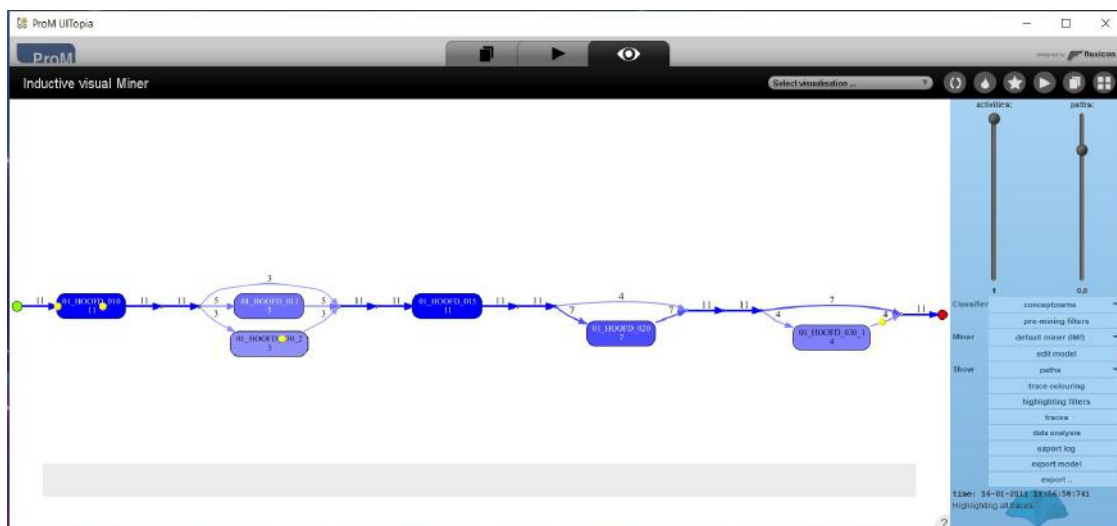
διαδικασίες και θα καταλήξουμε ποιος δήμος έχει τους λιγότερο αποδοτικούς υπαλλήλους. Μείζων θέμα που απαιτεί ιδιαίτερη προσοχή είναι τα cases τα οποία έχουν το μικρότερο αριθμό από ενεπ, πράγμα που σημαίνει ότι κάποιοι υπάλληλοι (οργανισμοί) δεν είναι όσο αποδοτικοί θα περιμέναμε.

1ο αρχείο (ελάχιστο μήκος 2 μέχρι 5)

Υποθέσεις: 11
 Δραστηριότητες: 41
 Υπάλληλοι: 5



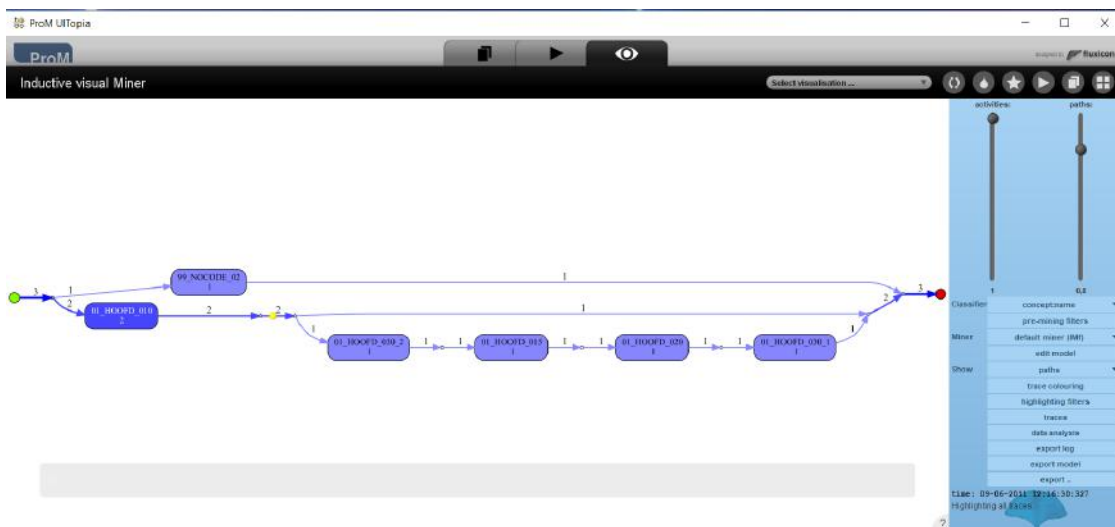
Inductive



Εικόνα 66: Dashboard και Process model 1ου αρχείου για ίχνη με λιγότερες από 5 ενέργειες

2ο αρχείο(ελάχιστο μήκος 1 μέχρι 5)

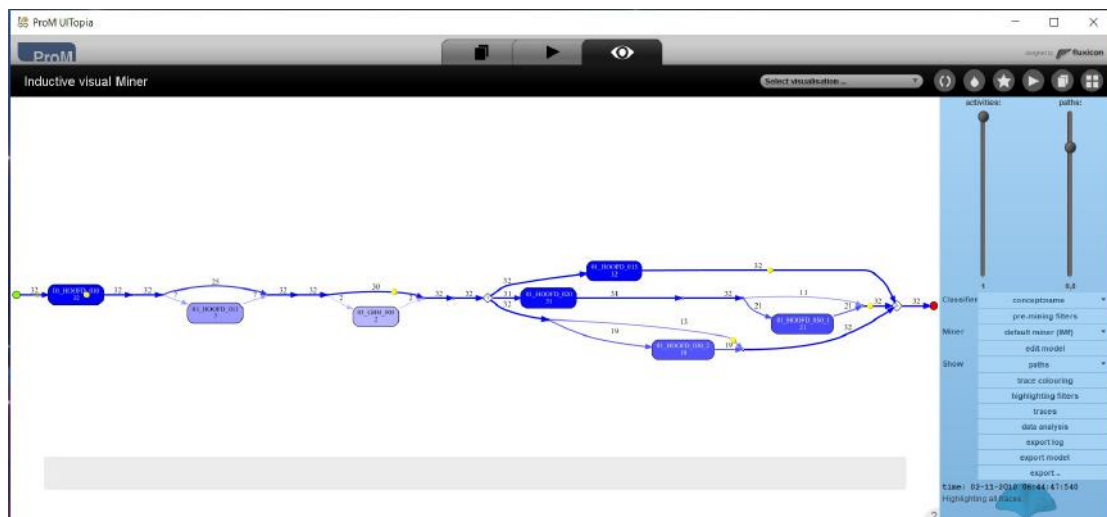
Υποθέσεις:3
 Δραστηριότητες:7
 Υπάλληλοι:3



Εικόνα 67: Dashboard και Process model 2ου αρχείου για ίχνη με λιγότερες από 5 ενέργειες

3ο αρχείο(ελάχιστο μήκος 3 μέχρι 5)

Υποθέσεις:32
 Δραστηριότητες:148
 Υπάλληλοι:5

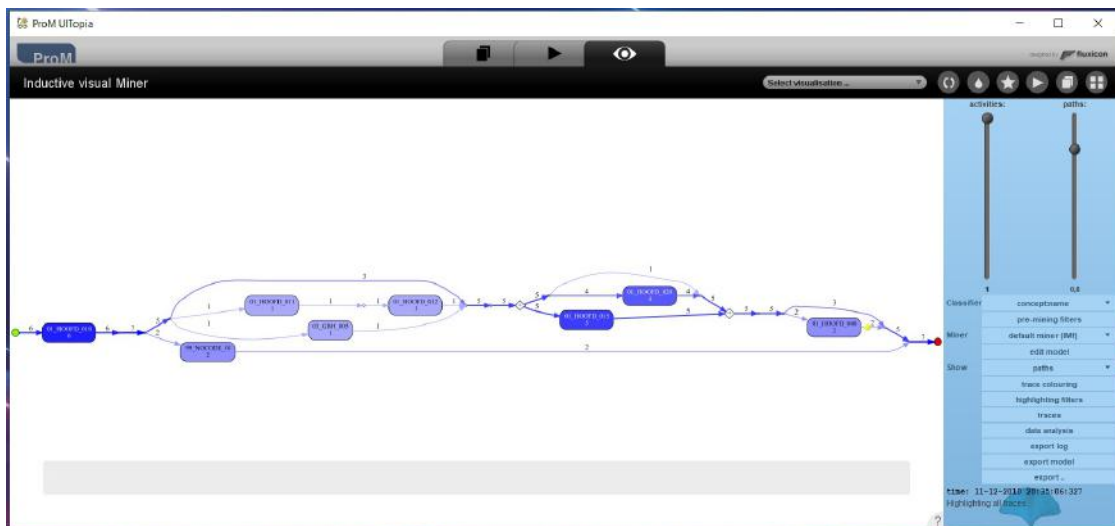


Εικόνα 68: Dashboard και Process model 3ου αρχείου για ίχνη με λιγότερες από 5 ενέργειες

4ο αρχείο(ελάχιστο μήκος 2 μέχρι 5)

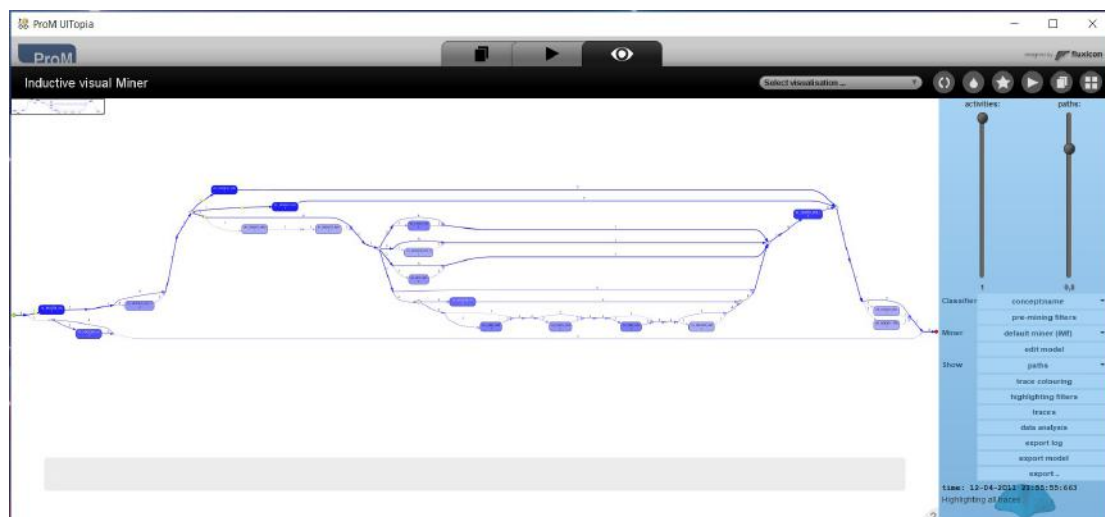


Υποθέσεις:7
 Δραστηριότητες:22
 Υπάλληλοι:3



Εικόνα 69: Dashboard και Process model 4ου αρχείου για ίχνη με λιγότερες από 5 ενέργειες

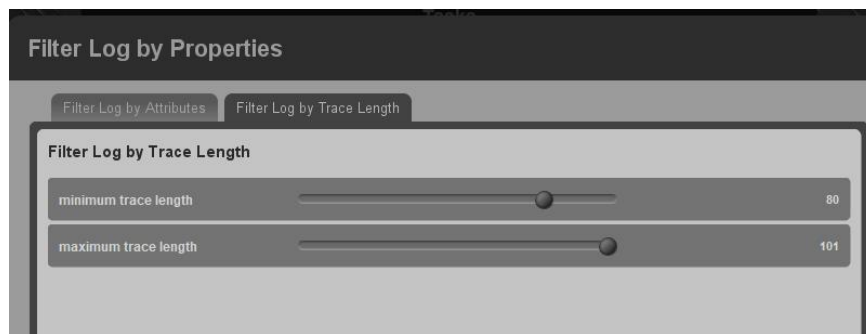
5ο αρχείο(ελάχιστο μήκος 5 μέχρι 10)
 Υποθέσεις:8
 Δραστηριότητες:58
 Υπάλληλοι:6



Εικόνα 70: Dashboard και Process model 5ου αρχείου για ίχνη με λιγότερες από 5 ενέργειες

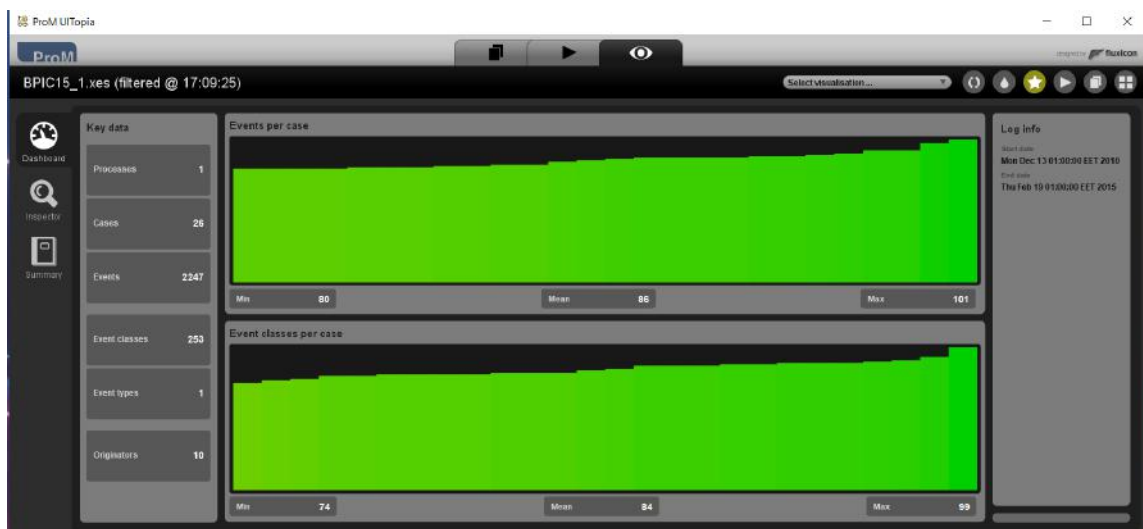
Παρατηρώντας τα αποτελέσματα βλέπουμε ότι το 2ο αρχείο είχε ελάχιστα cases με αριθμό event από 1 έως 5, το οποίο το καθιστά το δήμο με την υψηλότερη απόδοση και παραγωγικότητα μέχρι στιγμής. Στη συνέχεια ο 1ος και 4ος δήμος είχαν την αμέσως καλύτερη απόδοση με cases με λίγα events. Από την άλλη πλευρά ο 3ος δήμος είχε απογοητευτική απόδοση αυτήν την φορά, με 32 cases να έχουν από 3 έως 5 event (δεν υπήρχε case με λιγότερα από 3 event βέβαια στο δήμο). Αξιόλογο σημείο φαίνεται να είναι το 5ο αρχείο, το οποίο δεν είχε case με κάτω από 5 event και για αυτό η μέτρηση αφορούσε cases με 5 έως 10 event, πράγμα που το καθιστά έναν από τους πιο παραγωγικούς δήμους, με αρκετά πολυάσχολους φορείς και υπαλλήλου. Ο 2ος δήμος που έχει τις λιγότερες διαδικασίες παράγει, όπως αναμέναμε, ένα απλοϊκό μοντέλο χωρίς βρόγχους και επαναλήψεις. [23]

Στην επόμενη μέτρηση επιδιώκουμε με χρήση του ίδιου φίλτρου να δούμε την άλλη όψη του νομίσματος. Πιο συγκεκριμένα πριν είδαμε τους δήμους με υπαλλήλους που εκτελούν λίγες διαδικασίες, ενώ από την άλλη πλευρά τώρα θα δούμε τους δήμους που έχουν τους πιο πολυάσχολους υπαλλήλους. Θα εφαρμόσουμε ξανά το plug in **filter log by attributes** και ειδικά το **filter log by trace length**, θέτοντας κατώφλι την τιμή 80. Με αυτόν τον τρόπο θα μας εμφανιστούν τα αποτελέσματα για τα ίχνη με πάνω από 80 διαδικασίες εκτελεσμένες. Σε αυτήν την περίπτωση θέλουμε να δούμε μεγάλα ποσοστά, πράγμα που θα σημαίνει ότι ο εκάστοτε δήμος θα έχει αρκετά αποδοτικούς υπαλλήλους. Σε αυτήν την δοκιμή αφού παραχθούν τα φιλτραρισμένα αρχεία θα χρησιμοποιήσουμε την τεχνική οπτικοποίησης που είχαμε προαναφέρει το Dotted Chart διάγραμμα. Με σκοπό αφού δούμε τα στατιστικά των υπαλλήλων με πάνω από 80 εργασίες στο ενεργητικό τους, να παρατηρήσουμε πως συμπεριφέρονται σε βάθος 5ετίας και τι μοτίβα αναπτύσσουν.



Εικόνα 71: Filter Log by trace length

1ο αρχείο

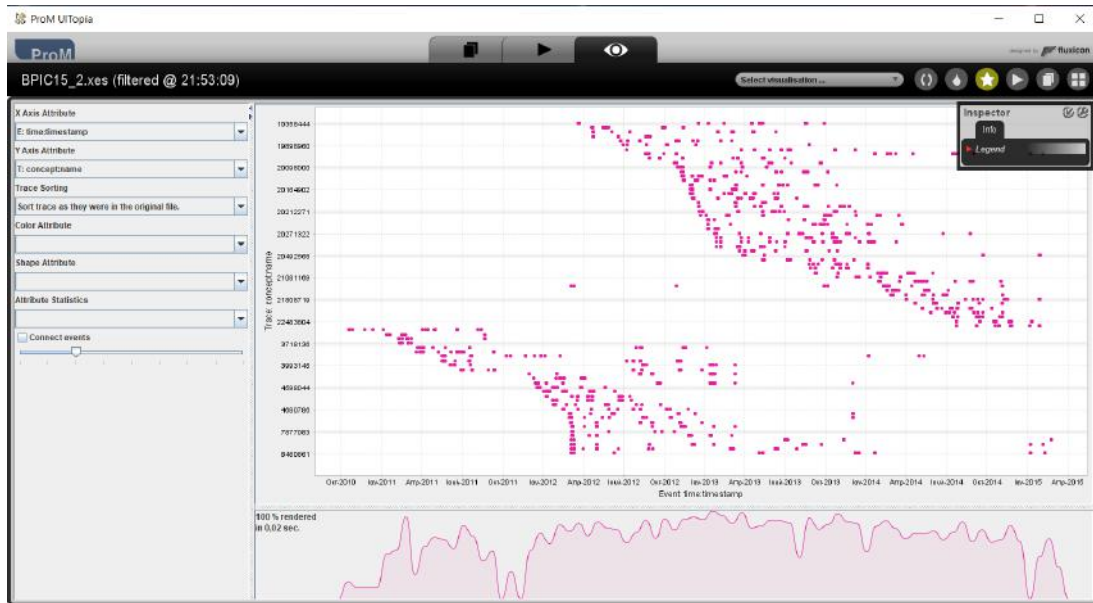




Εικόνα 72: Dashboard και Dotted Chart 1ου αρχείου για ίχνη με περισσότερες από 80 ενέργειες

2ο αρχείο

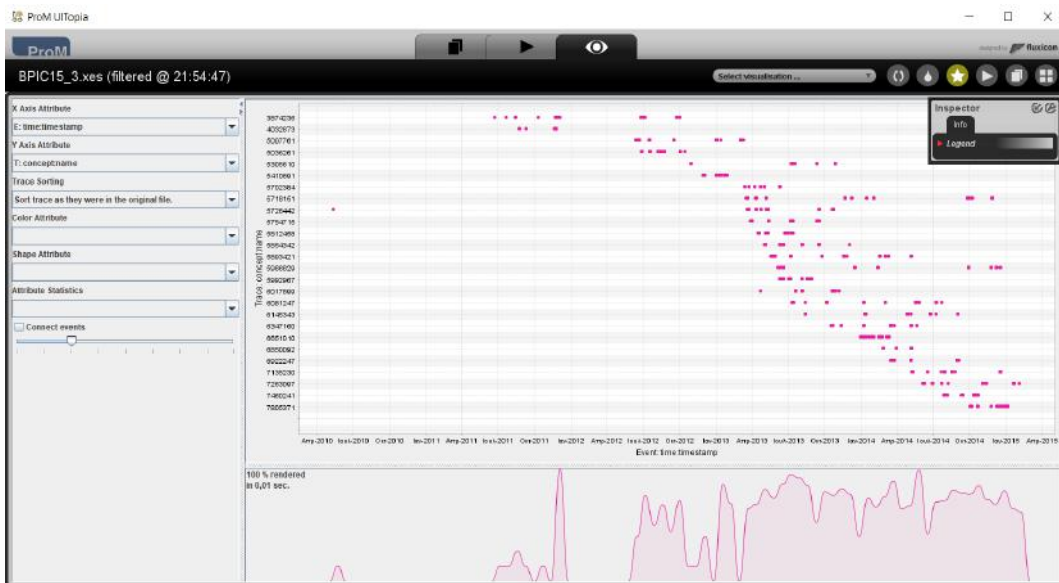




Εικόνα 73: Dashboard και Dotted Chart 2ου αρχείου για ίχνη με περισσότερες από 80 ενέργειες

3ο αρχείο

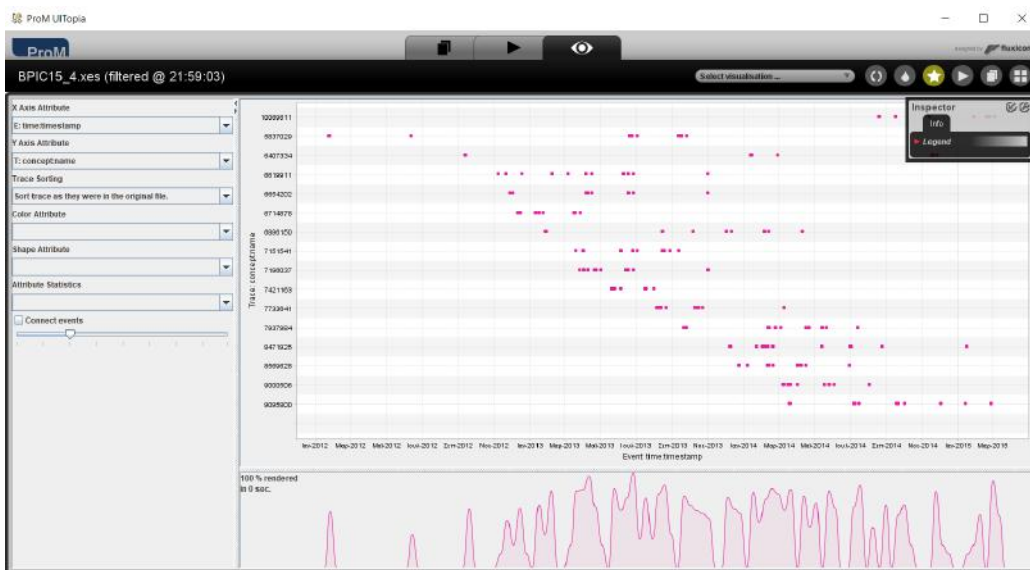




Εικόνα 74: Dashboard και Dotted Chart 3ου αρχείου για ίχνη με περισσότερες από 80 ενέργειες

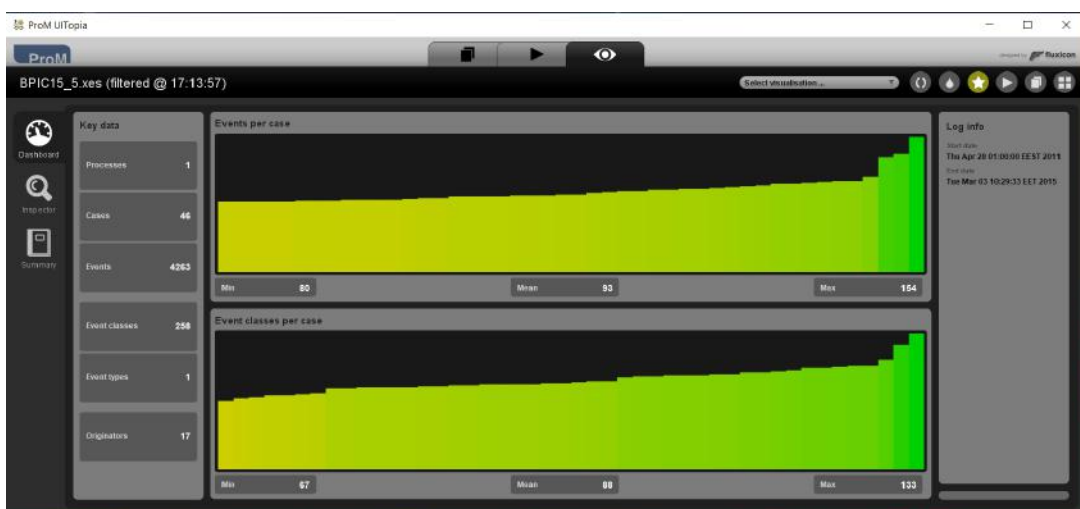
4ο αρχείο





Εικόνα 75: Dashboard και Dotted Chart 4ου αρχείου για ίχνη με περισσότερες από 80 ενέργειες

5ο αρχείο





Εικόνα 76: Dashboard και Dotted Chart 5ου αρχείου για ίχνη με περισσότερες από 80 ενέργειες

Συνεχίζοντας τις δοκιμές και σε αυτό το plug in έχουμε παρόμοια αποτελέσματα,διότι ο 2ος και ο 5ος δήμος έχουν τις περισσότερες υποθέσεις με πάνω από 80 event,ενώ χαμηλότερα σε απόδοση συνεχίζουν να είναι οι άλλοι 3 δήμοι.

Πιο συγκεκριμένα έχουμε:

- 1ος δήμος → cases 26,events 2247
- 2ος δήμος → cases 76,events 6922
- 3ος δήμος → cases 26,events 2313
- 4ος δήμος → cases 16,events 1384
- 5ος δήμος → cases 46,events 4263

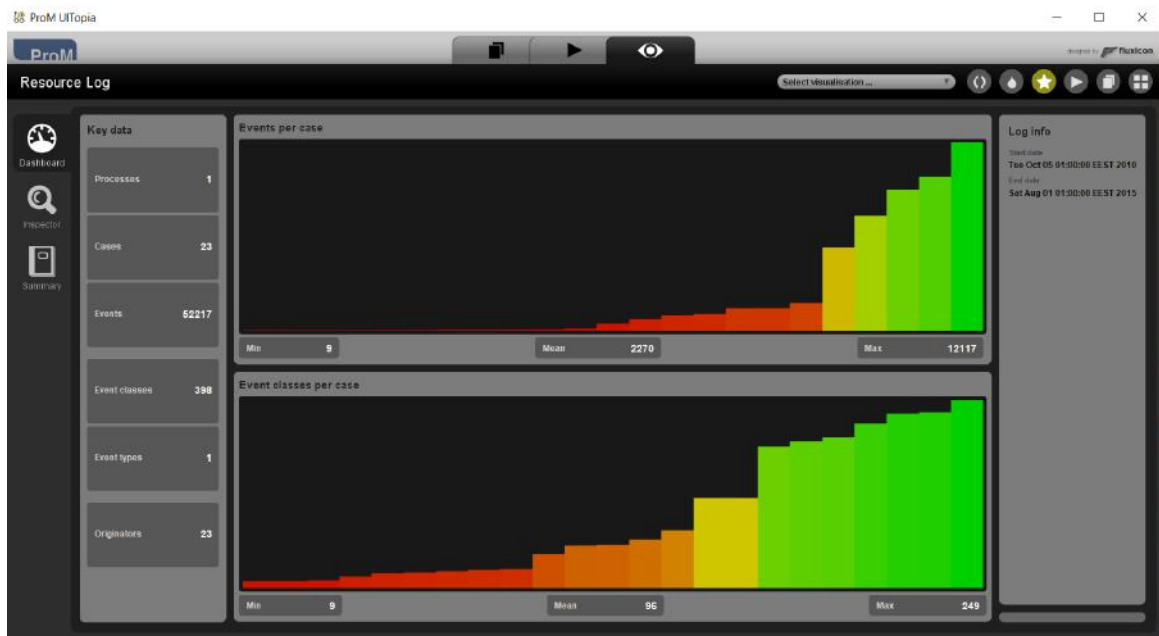
Βλέπουμε την απρόσμενα καλή απόδοση του 2ου δήμου και την αρκετά αξιόλογη απόδοση του 5ου δήμου.Στην αντίθετη πλευρά σε χαμηλά νούμερα κυμαίνονται οι υπόλοιποι δήμοι,με την απόδοση του 2ου δήμου να φτάνει μέχρι και 6 φορές πάνω από αυτή του 4ου.Ακόμα μεγάλο ενδιαφέρον συγκεντρώνει ότι ο 5ος δήμος είχε trace με μέχρι 154 events και ο 2ος είχε μέχρι 132 events.Η εφαρμογή του φίλτρου έγινε με πάνω από 80 event μέχρι το ανώτατο όριο του κάθε δήμου.Με τα παρακάτω δεδομένα βλέπουμε ότι ο 2ος δήμος είχε μέχρι και 7 φορές παραπάνω event από τα αντίστοιχα trace του 4ου δήμου,που είχε την χαμηλότερη απόδοση.Παρατηρούμε λοιπόν στο

σημείο αυτό την υψηλή παραγωγικότητα των υπαλλήλων του δήμου 2 και 5,έχοντας αρκετά μεγάλο αριθμό εργαζομένων με πάνω από 80 εκτελεσμένες διαδικασίες,πράγμα που σημαίνει ότι σε αυτούς τους δήμους υπάρχει υψηλή απόδοση και παραγωγή,

Τα παραπάνω στοιχεία έρχονται να επιβεβαιώσουν και τα dotted chart διαγράμματα τα οποία μας δείχνουν στον κάθετο άξονα τα cases και στον οριζόντιο άξονα τον χρόνο,δείχνοντας τα activities που εκτελεί ο κάθε οργανισμός σε συνάρτηση με το χρόνο.Βλέπουμε στα αρχεία 2 και 5 πυκνότερα διαγράμματα που περιέχουν πολλές περισσότερες εργασίες σε σχέση με τους υπόλοιπους δήμους.Ακόμα οι κυματομορφές κάτω από τα διαγράμματα δείχνουν τα πολυπληθή cases των δήμων 2 και 5,οι οποίες παραμένουν σε υψηλό δείκτη όλη την διάρκεια της 5ετίας,σε αντίθεση με τους υπόλοιπους δήμους που φτάνουν σε ένα peak της απόδοσης τους και μετά υποχωρούν.Αξιοσημείωτο σημείο φαίνεται να είναι ότι σε όλα τα αρχεία η κορύφωση των εργασιών σημειώνεται περίπου το πρώτο 6μηνο του 2013,όπου εκεί φαίνεται να ήταν η πιο παραγωγική περίοδος για τους δήμους.

Καθώς πριν ασχοληθήκαμε με τους εργαζομένους των δήμων και τις διαδικασίες που εκτελούν,θα ήθελα στην επόμενη μέτρηση να ασχοληθώ με την κατανομή των εργαζομένων στους δήμους.Με αυτόν τον τρόπο θα δούμε πραγματικά τι συμβαίνει στους δήμους και πως διαμοιράζονται οι εργασίες σε κάθε υπεύθυνο.Αυτή η μέτρηση για να έχουμε μία καλύτερη εικόνα των διεργασιών και της εξόρυξης αυτών,έγινε με την χρήση του plug in Generate log from org:resource perspective στο prom tools.Σκοπός αυτού του φίλτρου ήταν να μας δείξει για κάθε αρχείο δήμου πόσοι org:resources υπάρχουν,για πόσα events είναι υπεύθυνοι και γενικά να δούμε αναλυτικότερα πόσες διεργασίες αναλαμβάνει ο κάθε οργανισμός τόσο σε μέσο όρο αλλά και τους οργανισμούς με τις περισσότερες και λιγότερες διεργασίες αντίστοιχα.Με αυτόν τον τρόπο θα εξάγουμε συμπεράσματα για τους πιο παραγωγικούς οργανισμούς,καθώς δεν αρκεί να εκτελούνται πολλές εργασίες,θα πρέπει και το φόρτο αυτών να μοιράζεται σωστά σε όλους.Ούτε ωφελεί κανέναν ,πόσο μάλλον τον ίδιο το δήμο, να έχει πάρα πολλούς υπαλλήλους χωρίς να εκτελούνται αντίστοιχα πολλές εργασίες,ούτε και να εκτελεί πολλές εργασίες με πάρα πολλούς εργαζομένους.
Αναλυτικότερα έχουμε τα εξής στοιχεία:

1ο αρχείο



Εικόνα 77: Generate log from org:resource perspective για τον 1ο δήμο

2ο αρχείο



Εικόνα 78: Generate log from org:resource perspective για τον 2ο δήμο

3ο αρχείο



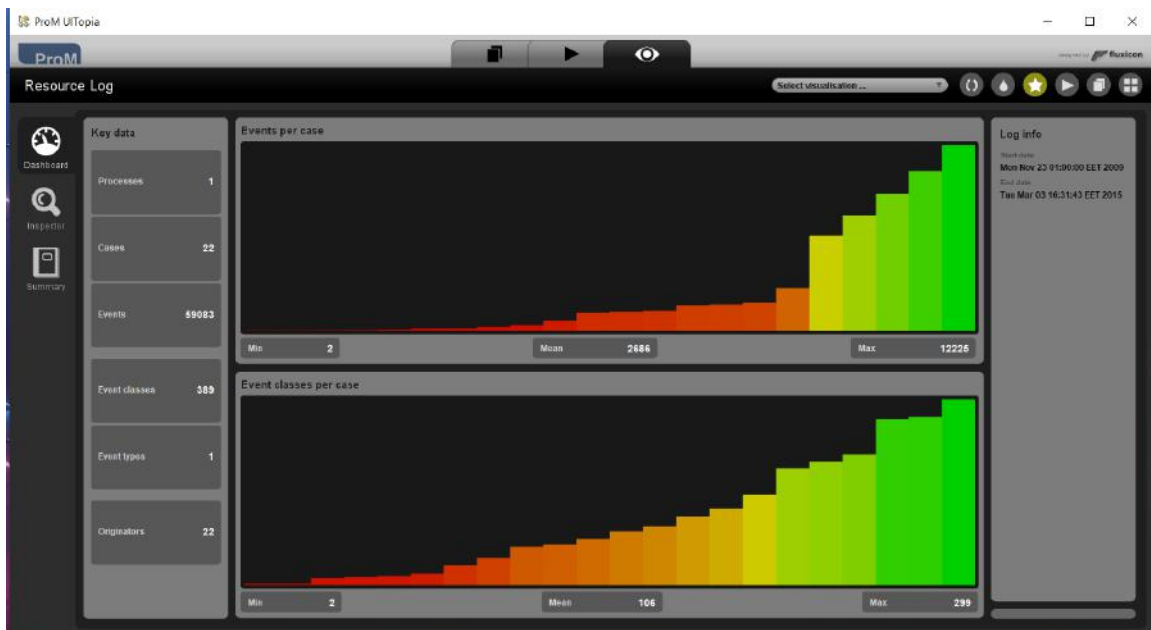
Εικόνα 79: Generate log from org:resource perspective για τον 3ο δήμο

4ο αρχείο



Εικόνα 80: Generate log from org:resource perspective για τον 4ο δήμο

5ο αρχείο



Εικόνα 81: Generate log from org:resource perspective για τον 5ο δήμο

Όπως φαίνεται στα παραπάνω αποτελέσματα έχω παραθέσει για αρχή την καρτέλα log visualizer από το prom tools, όπου μπορούμε να δούμε τον αριθμό των resources, πόσα events αναλαμβάνει, κατά μέσο όρο, το κάθε resource και ποίο resource διεκπεραιώνει τα λιγότερα και περισσότερα event.

1ος δήμος → 23
 2ος δήμος → 11
 3ος δήμος → 14
 4ος δήμος → 10
 5ος δήμος → 22

Με τα παραπάνω βλέπουμε τον 1ο και 5ο δήμο να έχουν τα περισσότερα resources όπως είναι λογικό, γιατί έχουν και πολλά event προς εκτέλεση, συγκεκριμένα άνω των 50.000. Ειδικά ο 5ος δήμος μας έχει συνηθίσει σε υψηλά ποσοστά παραγωγικότητας και απόδοσης. Από την αντίπερα όχθη ο 3ος δήμος με σχεδόν 60.000 events βλέπουμε ότι απασχολεί μόλις 14 υπηρεσίες για να τα εκτελέσει, πράγμα αξιοθαύμαστο για τον δήμο. Ενώ ο 2ος και 4ος δήμος έχοντας τα λιγότερα event, κάτω των 50.000, αναμενόμενα απασχολούν τους λιγότερους οργανισμούς.

Βέβαια είναι αναγκαίο να σημειωθεί ότι σε πολλές περιπτώσεις οργανισμοί/υπάλληλοι συνεργάζονται και με παραπάνω από 1 δήμο την φορά, οπότε δεν μπορούμε να μπορούμε να βγάλουμε τόσο αξιόπιστα αποτελέσματα.

1ος δήμος → 2270 mean events, 12117 max events
 2ος δήμος → 4032 mean events, 11479 max events
 3ος δήμος → 4263 mean events, 14620 max events

4ος δήμος → 4729 mean events, 15748 max events
 5ος δήμος → 2686 mean events, 12225 max events

Παρόλα αυτά πρέπει να σχολιάσουμε ότι οι εργαζόμενοι στους δήμους 2,3 και 4 έχουν κατά μέσο όρο πολλοί περισσότερες εργασίες ξεπερνώντας τις 4.000,σε αντίθεση με τους δήμους 1 και 5 που βρίσκονται περίπου στις 2.500 εργασίες.Σημειώνοντας επίσης ότι ο 4ος δήμος έχει οργανισμό με πάνω από 15.000 event και ότι ο οργανισμός με τις λιγότερες εργασίες στον 2ο δήμο,εκτελεί 19 activities,την στιγμή που στους υπόλοιπους δήμους υπάρχουν υπάλληλοι με 2 ή 3 δραστηριότητες.

Επίσης στη συνέχεια θα ήθελα να παρουσιάσω και τους εργαζομένους που βρίσκονται,όπως είχα προαναφέρει,σε παραπάνω από ένα δήμους και πως αποδίδουν σε αυτό.

Αναλυτικότερα έχουμε τα εξής στοιχεία:

Τον οργανισμό/υπάλληλο (org:resource) με κωδικό

- 6 με εργασίες στον 1ο,3ο και 4ο δήμο.

Στον 1ο δήμο έχει 26 εκτελεσμένες εργασίες,ενώ στον 3ο και 4ο έχει 2 και 3 αντίστοιχα όπως προαναφέραμε,καθώς είναι ο οργανισμός με τις λιγότερες διεργασίες στους δήμους αυτούς.

- 560429 με εργασίες στον 2ο και 5ο δήμο.

Στον 2ο έχει 19 εργασίες και στον 5ο 7.590 εργασίες.

- 560530 με εργασίες στον 2ο και 5ο δήμο.

Στον 2ο έχει 11.479 εργασίες και στον 5ο 683 εργασίες.

- 560532 με εργασίες στον 2ο και 5ο δήμο.

Στον 2ο έχει 10.080 εργασίες και στον 5ο 1.317 εργασίες.

- 560598 με εργασίες στον 2ο και 5ο δήμο.

Στον 2ο έχει 183 εργασίες και στον 5ο 1.737 εργασίες.

- 560752 με εργασίες στον 4ο και 5ο δήμο.

Στον 4ο έχει 11.948 εργασίες και στον 5ο 1.676 εργασίες.

- 560849 με εργασίες στον 4ο και 5ο δήμο.

Στον 4ο έχει 764 εργασίες και στον 5ο 154 εργασίες.

Βλέπουμε παραπάνω ότι στις 6 από τις 7 περιπτώσεις οι υπάλληλοι αυτοί, ανήκουν στον 5ο δήμο.Ενώ στις 4 από αυτές οι υπάλληλοι είναι μεταξύ του 2ου και 5ου δήμου και στις άλλες 2 μεταξύ του 4ου και 5ου δήμου.Αυτές οι ενδείξεις έρχονται να συμπληρώσουν την άποψη για την υψηλή παραγωγικότητα του 2ου και 5ου δήμου,αποδεικνύοντας ότι με την κατάλληλη συνεργασία

μπορούν να διεκπαιρωθούν τα ζητήματα των υπηρεσιών και μάλιστα σε μικρότερο χρονικό διάστημα.

Με βάση τις παραπάνω μετρήσεις που εκτελέστηκαν μπορούμε να καταλήξουμε σε κάποια συμπεράσματα για την λειτουργία των δήμων. Για παράδειγμα παρατηρούμε ότι ο δήμος νούμερο 2 και νούμερο 5 έχουν την υψηλότερη παραγωγικότητα, με αρκετά μεγάλο μέσο όρο εργασιών ανά ημέρα, γεγονός που τους κάνει να εκτελούν πολλές διεργασίες καθημερινά και ταχύτατα. Ο 5ος δήμος στο διάστημα της 4ετίας εκτελεί τις περισσότερες διαδικασίες. Σε αντίθεση ο χαμηλός ρυθμός εκτέλεσης εργασιών, κάνει τους υπόλοιπους δήμους να καθυστερούν στην διεκπεραίωση ζητημάτων. Επίσης ο 2ος και 5ος δήμος έχουν και τους πιο παραγωγικούς υπαλλήλους, που εκτελούν πολλές παραπάνω εργασίες, ενώ ταυτόχρονα καταφέρνουν να συνεργάζονται μεταξύ τους για την καλύτερη εκπόνηση εργασιών. Συγκεκριμένα ο 2ος δήμος είχε μόλις δύο “απομονωμένους” υπαλλήλους και όλοι οι υπόλοιποι, συνεργάζονταν με άλλους για την εκτέλεση διαδικασιών. Το ίδιο συμβαίνει και με τον 5ο δήμο όπου όλοι οι εργαζόμενοι συνεισφέρουν και στους άλλους δήμους, ακόμα και στον 4ο δήμο, που εξακολουθεί να έχει μειωμένη απόδοση. Ειδικά αν αναλογιστούμε ότι η ημερομηνία έναρξης εργασιών του 2ου δήμου είναι 6 μήνες αργότερα από τον 4ο δήμο, που ξεκινάει πρώτος, και παρόλα αυτά εκτελεί αρκετές παραπάνω εργασίες.

Από την άλλη όψη του νομίσματος οι δήμοι 1, 3 και 4 έχουν εργαζομένους που δεν συνεργάζονται με άλλους δήμους, πράγμα που από ότι φαίνεται μειώνει την αποδοτικότητα τους. Έχουν χαμηλό μέσο όρο εκτέλεσης εργασιών ανά εβδομάδα, ενώ ειδικά ο 1ος δήμος έχει τους περισσότερους υπαλλήλους και καταφέρει να εκτελούν τις λιγότερες εργασίες στο μέσο όρο. Τέλος έρχεται να προστεθεί και το πλήθος των εργαζομένων με πάνω από έναν αριθμό εργασιών (κατώφλι στην περίπτωση μας οι 80 διαδικασίες) όπου οι εργαζόμενοι του 2ου και 5ου δήμου υπερτέρουν ξανά, δείχνοντας ότι το σύνολο αυτών, και όχι μόνο μία μερίδα, είναι πολύ πιο αποδοτικοί και αναλαμβάνουν με σωστή κατανομή εργασίες. Στην αντίστοιχη έρευνα για τους εργαζομένους με μικρό αριθμό εργασιών (κατώφλι στην περίπτωση μας οι 5 εργασίες) βλέπουμε τους δήμους 1 και 3 να έχουν μεγάλο πλήθος εργαζομένων με πολλή λίγες διαδικασίες στο ενεργητικό τους. Και σε αυτήν την περίπτωση οι δήμοι 2 και 5 κατάφεραν να υπερισχύουν με τον 2ο δήμο να έχει λίγες υποθέσεις με κάτω από 5 διαδικασίες, ενώ ο 5ος δήμος δεν είχε υπόθεση με κάτω από 5 εργασίες.

ΚΕΦΑΛΑΙΟ 5 ΣΥΜΠΕΡΑΣΜΑ

Με όλα τα παραπάνω αποτελέσματα και στοιχεία, βλέπουμε ότι ο αριθμός των διαδικασιών που εκτελούνται καθημερινά στο επαγγελματικό επίπεδο του κάθε ανθρώπου, είναι τεράστιος και επηρεάζει κατά πολύ την βιωσιμότητα και απόδοση της επιχείρησης και των ατόμων που βρίσκονται σε αυτή. Η τεχνική της εξόρυξης διαδικασιών με τα εργαλεία της μπορεί να αναλύσει τις διαδικασίες που λαμβάνουν χώρα, να μελετήσει την δομή και τον τρόπο λειτουργίας τους και να βγάλει συμπεράσματα για αυτές.

Τα αρχεία καταγραφής συμβάντων από τους δήμους και γενικά των επιχειρήσεων μας προσφέρουν μία πληθώρα πληροφοριών προς μελέτη και επεξεργασία. Αναλύοντας τα με χρήση τεχνικών εξόρυξης, μπορούμε να δούμε τον τρόπο εκτέλεσης διαδικασιών, αν σε κάποια περίπτωση έχουμε χαμηλή απόδοση από κάποιο δήμο ή εργαζόμενο και γιατί συγκεκριμένα, αν τα μοντέλα διαδικασιών συμφωνούν με τα δεδομένα των αρχείων και ίσως τις περιπτώσεις που μπορεί να υπάρχει και αδικαιολόγητη σπατάλη πόρων για τις εργασίες.

Με τα εργαλεία της εξόρυξης διαδικασιών που εφαρμόσαμε σε αυτήν την διπλωματική, βγάζουμε χρήσιμα συμπεράσματα για τα ερωτήματα που τέθηκαν στην αρχή της παρουσίασης. Όπως ποία/ες είναι οι πιο συνηθισμένες διαδικασίες στους δήμους, όπου αποδεικνύεται ότι η εργασία με κωδικό 01_HOOFD_100 (register submission date request) είναι στους 4 από τους 5 δήμους, η πιο συχνά εκτελέσιμη εργασία. Ακόμα στο ερώτημα για τους πιο παραγωγικούς εργαζομένους βλέπουμε ότι εκείνοι στον 2ο και 5ο δήμο εκτελούν αρκετές εργασίες σε σχέση με τους υπόλοιπους, εξασφαλίζοντας στους αντίστοιχους δήμους υψηλή παραγωγικότητα. Ιδιαίτερα στον 2ο δήμο που διαθέτει και λίγους υπαλλήλους. Επισημαίνοντας με βάση τα αποτελέσματα, ότι πολλοί εργαζόμενοι στο 2ο δήμο εκτελούν εργασίες και στον 5ο δήμο και ανάποδα. Δείχνοντας ότι σε περιπτώσεις που υπάρχει συνεργασία και αλληλοβοήθεια, επιτυγχάνονται καλύτερα αποτελέσματα. Ωστόσο από την άλλη πλευρά βλέπουμε τους υπόλοιπους 3 δήμους με χαμηλότερη παραγωγή, σε σχέση και με το ανθρώπινο δυναμικό που διαθέτουν και την ελάχιστη συνεργασία που δείχνουν να έχουν μεταξύ τους. Το συμπέρασμα αυτό επιβεβαιώνει ο μεγάλος χρόνος, που απαιτείται για την εκπόνηση εργασιών στους δήμους 1, 3 και 4, καθώς και ο χαμηλός μέσος όρος που κατέχουν. Στις μετρήσεις βλέπαμε εξ αρχής τα στατιστικά στοιχεία των δήμων και σε συμφωνία με τα οπτικά αποτελέσματα (διαγράμματα, μοντέλα διαδικασιών) που παρήχθησαν στην πορεία, μας επιβεβαιώνουν τα συμπεράσματα που βγάλαμε.

Η εξόρυξη διαδικασιών για την ανάλυση και αξιολόγηση των επιχειρηματικών διαδικασιών στην τοπική αυτοδιοίκηση, είναι ένα μείζον ζήτημα με ευρεία κλίμακα αναζήτησης, που μπορεί να μας βοηθήσει να κατανοήσουμε καλύτερα τον τρόπο εκτέλεσης διαδικασιών και να συνεισφέρουμε στην βελτίωση αυτών, με σκοπό να προσφέρουμε μία ικανοποιητική απόδοση για τους δήμους και τους ανθρώπους.

ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ

- [1] [Online] <https://el.wikipedia.org/wiki/%CE%95%CF%80%CE%B9%CF%87%CE%B5%CE%AF%CF%81%CE%B7%CF%83%CE%B7>
- [2] [Online] https://el.wikipedia.org/wiki/%CE%9F%CF%81%CE%B3%CE%B1%CE%BD%CE%B9%CF%83%CE%BC%CE%BF%CE%AF_%CE%A4%CE%BF%CF%80%CE%B9%CE%BA%CE%AE%CF%82_%CE%91%CF%85%CF%84%CE%BF%CE%B4%CE%B9%CE%BF%CE%AF%CE%BA%CE%B7%CF%83%CE%B7%CF%82
- [3] [Online] https://en.wikipedia.org/wiki/Business_intelligence
- [4] [Online] [https://www.gartner.com/en/information-technology/glossary/business-process-management-bpm#:~:text=Business%20process%20management%20\(BPM\)%20is,support%20of%20a%20business%20strategy](https://www.gartner.com/en/information-technology/glossary/business-process-management-bpm#:~:text=Business%20process%20management%20(BPM)%20is,support%20of%20a%20business%20strategy)
- [5] [Online] [https://www.investopedia.com/terms/b/business-intelligence-bi.asp#:~:text=Business%20intelligence%20\(BI\)%20refers%20to,performance%20benchmarking%2C%20and%20descriptive%20analytics](https://www.investopedia.com/terms/b/business-intelligence-bi.asp#:~:text=Business%20intelligence%20(BI)%20refers%20to,performance%20benchmarking%2C%20and%20descriptive%20analytics).
- [6] [Online] https://en.wikipedia.org/wiki/Data_mining
- [7] [online] <http://www.processmining.org/introduction.html>
- [8] A Study of Process Mining-based Business Process Innovation, Sungbum Parka, Young Sikkang
- [9] Process Mining plays an essential role in Digital Transformation, Pedro Robledo
BPM process management expert
- [10] [Online] <https://fluxicon.com/blog/2010/10/prom-tips-mining-algorithm/>
https://en.wikipedia.org/wiki/Alpha_algorithm
- [11] [Online] <http://www.processmining.org/event-data.html>
- [12] Towards Cross-Organizational Process Mining in Collections of Process Models and Their Executions, J.C.A.M. Buijs, Boudewijn F. van Dongen, Wil M.P. van der Aalst
- [13] [Online] <https://www.promtools.org/doku.php>
- [14] [Online] <https://eventlogxp.com/>
- [15] [Online] <https://www.futurelearn.com/info/courses/process-mining/0/steps/15636>
- [16] Heuristic mining: Adaptive process simplification in education, Publisher: IEEE
- [17] [Online] <https://www.futurelearn.com/info/courses/process-mining/0/steps/15645>
- [18] Discovery and analysis of the Dutch permitting process, Liese Blevi and Peter Van den Spiegel

- [19] Benchmarking of Municipalities with Process Mining Techniques Reveals Opportunities for Improvement ,Ube van der Ham
- [20] Handling of Building Permit Applications in The Netherlands: A Multi-Dimensional Analysis, P.M. Dixit , B.F.A. Hompes, N. Tax, S.J. van Zelst
- [21] [Online] <https://research.aimultiple.com/process-mining-algorithms/>
- [22] An Exploration and Analysis of The Building Permit Application Process in Five Dutch Municipalities,Niels Martin, Gert Janssenswillen, Toon Jouck, Marijke Swennen, Mehrnush Hosseinpour, Farahnaz Masoumigoudarzi
- [23] Process Mining,Wil Van Der Aalst