



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
UNIVERSITY OF PIRAEUS

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

ΤΜΗΜΑ ΨΗΦΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
«ΠΛΗΡΟΦΟΡΙΑΚΑ ΣΥΣΤΗΜΑ & ΥΠΗΡΕΣΙΕΣ»

«ΜΕΓΑΛΑ ΔΕΔΟΜΕΝΑ ΚΑΙ ΑΝΑΛΥΤΙΚΗ»

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ ΔΙΠΛΩΜΑΤΟΣ ΕΙΔΙΚΕΥΣΗΣ

**Πειραματική Αποτίμηση Αλγορίθμων Ανάλυσης
Διεργασιών**

ΧΡΗΣΤΟΣ ΣΤΑΜΑΤΑΚΗΣ

ΑΘΗΝΑ

ΦΕΒΡΟΥΑΡΙΟΣ 2020

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Πειραματική Αποτίμηση Αλγορίθμων Ανάλυσης Διεργασιών

ΧΡΗΣΤΟΣ ΣΤΑΜΑΤΑΚΗΣ

A.M.: ME1730

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ:

Δουλκερίδης Χρήστος, Επίκουρος Καθηγητής

ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ

Τελέλης Ορέστης, Επίκουρος Καθηγητής

Φιλιππάκης Μιχαήλ, Αναπληρωτής Καθηγητής

ΗΜΕΡΟΜΗΝΙΑ ΕΞΕΤΑΣΗΣ

06/03/2020

Περίληψη

Σκοπός της παρούσας διπλωματικής εργασίας είναι η εφαρμογή, μελέτη και σύγκριση τεσσάρων διαφορετικών αλγορίθμων ανάλυσης διεργασιών (process mining). Οι αλγόριθμοι που θα μελετηθούν και θα εφαρμοστούν είναι οι εξής: Alpha Miner, Inductive Miner, Evolutionary Tree Miner και Heuristic Miner. Σκοπός της εργασίας είναι η συγκριτική αποτίμηση των ανωτέρω αλγορίθμων, με έμφαση στην ποιότητα των παραγόμενων αποτελεσμάτων. Για τον προσδιορισμό της ποιότητας των αποτελεσμάτων που θα προκύψουν από τους αλγόριθμους ανάλυσης διεργασιών θα χρησιμοποιηθούν και θα καταγραφούν μετρικές ποιότητας όπως: Simplicity, Generalization, Precision και Fitness.

Για την εφαρμογή των αλγορίθμων χρησιμοποιήθηκαν συνθετικά δεδομένα που αναπαριστούν μοντέλα διεργασιών. Η παραγωγή των μοντέλων διεργασιών θα γίνει με την χρήση του λογισμικού PLG2 και η μελέτη των αλγορίθμων ανάλυσης διεργασιών με την χρήση του λογισμικού Prom.

Κατά την διάρκεια των πειραματισμών έλαβαν χώρα διάφορες παραμετροποιήσεις στα μοντέλα διεργασιών. Οι σχετικές παράμετροι θα είναι οι χρόνοι διάρκειας των δραστηριοτήτων των μοντέλων, οι επαναλήψεις εκτέλεσης του μοντέλου, οι εναλλαγές στην ποσότητα του θορύβου του μοντέλου αλλά και οι εναλλαγές στις σχέσεις μεταξύ των διεργασιών του μοντέλου.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Ανάλυση Διεργασιών.

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: PLG2, Prom, Αλγόριθμοι, Μοντέλα Διεργασιών, Εξόρυξη διεργασιών, Μετρικές Ποιότητας

Abstract

The purpose of this thesis is to apply, study and compare four different process analysis algorithms (process mining). The algorithms to be studied and implemented are: Alpha Miner, Inductive Miner, Evolutionary Tree Miner and Heuristic Miner. The purpose of this work is to compare the above algorithms with emphasis on the quality of the outputs. Process analysis algorithms will be used to record quality metrics such as: Simplicity, Generalization, Precision and Fitness.

In order to test the algorithms, it has been used artificial data which represents process models. The process models will be produced using PLG2 software and the study of process analysis algorithms using Prom software.

During the experiments, various customizations were made to the models. The relevant parameters will be the duration of the model activities, the iterations of the execution model, the variations in the amount of model noise, as well as the changes in the relationships between the model processes.

SUBJECT AREA: Process Mining.

KEYWORDS: PLG2, Prom, Algorithms, Process Models, Process Mining, Quality Metrics

Περιεχόμενα

1	ΕΙΣΑΓΩΓΗ.....	15
1.	PROCESS MINING	16
1.1	ΟΡΙΣΜΟΣ ΤΟΥ PROCESS MINING.....	16
1.2	ΕΦΑΡΜΟΓΕΣ ΤΟΥ PROCESS MINING	18
1.3	ΜΟΝΤΕΛΑ ΔΙΕΡΓΑΣΙΩΝ ΚΑΙ ΔΕΔΟΜΕΝΑ.....	19
2	ΠΕΡΙΓΡΑΦΗ PROCESS MINING SOFTWARE.....	26
2.1	ΛΟΓΙΣΜΙΚΟ ΔΗΜΙΟΥΡΓΙΑΣ ΤΥΧΑΙΩΝ ΔΙΕΡΓΑΣΙΩΝ PLG2	26
2.2	ΤΟ ΕΡΓΑΛΕΙΟ PROM	27
3	ΑΛΓΟΡΙΘΜΟΙ ΚΑΙ ΜΕΤΡΙΚΕΣ ΠΟΙΟΤΗΤΑΣ.....	29
3.1	ΑΛΓΟΡΙΘΜΟΙ.....	29
3.1.1	<i>Alpha Miner</i>	29
3.1.2	<i>Heuristic Miner</i>	31
3.1.3	<i>Evolutionary Tree Miner</i>	32
3.1.4	<i>Inductive Visual Miner</i>	33
3.2	ΜΕΤΡΙΚΕΣ ΠΟΙΟΤΗΤΑΣ.....	34
4	ΔΙΕΞΑΓΩΓΗ ΠΕΙΡΑΜΑΤΩΝ.....	38
4.1	ΕΙΣΑΓΩΓΗ ΣΤΗΝ ΔΙΑΔΙΚΑΣΙΑ ΔΙΕΞΑΓΩΓΗΣ ΠΕΙΡΑΜΑΤΩΝ	38
4.2	ΠΑΡΑΓΩΓΗ ΣΥΝΘΕΤΟΥ ΜΟΝΤΕΛΟΥ ΜΕ ΤΗΝ ΧΡΗΣΗ ΤΟΥ ΛΟΓΙΣΜΙΚΟΥ PLG2 39	
4.3	ΠΕΙΡΑΜΑ ΜΕ 500 ΕΚΤΕΛΕΣΕΙΣ.....	40
4.4	ΠΕΙΡΑΜΑ ΜΕ 1000 ΕΚΤΕΛΕΣΕΙΣ.....	51
4.5	ΠΕΙΡΑΜΑ ΜΕ 1500 ΕΚΤΕΛΕΣΕΙΣ.....	62
4.6	ΕΥΑΛΟΥΑΤΙΟΝ ΤΩΝ PROCESS MINING ΜΟΝΤΕΛΩΝ	72
4.7	ΠΙΝΑΚΕΣ ΜΕΤΡΙΚΩΝ ΠΟΙΟΤΗΤΑΣ	77
5	ΣΥΜΠΕΡΑΣΜΑΤΑ – ΜΕΛΛΟΝΤΙΚΕΣ ΒΕΛΤΙΩΣΕΙΣ	80
6	ΒΙΒΛΙΟΓΡΑΦΙΑ.....	82

Κατάλογος Εικόνων

ΕΙΚΟΝΑ 1 ΜΟΝΤΕΛΟ ΜΕ ΣΗΜΕΙΟΓΡΑΦΙΑ ΤΟΥ ΒΡΜΝ.....	19
ΕΙΚΟΝΑ 2 TRACES EVENT LOG ΑΠΟ ΕΚΤΕΛΕΣΗ ΒΡΜΝ (ΠΕΔΙΟ TRACE).....	21
ΕΙΚΟΝΑ 3 TRACES EVENT LOG ΑΠΟ ΕΚΤΕΛΕΣΗ ΒΡΜΝ (ΠΕΔΙΟ TIMESTAMP)	22
ΕΙΚΟΝΑ 4 PETRI NET	23
ΕΙΚΟΝΑ 5 PROCESS TREE	24
ΕΙΚΟΝΑ 6: ΠΑΡΑΓΩΓΗ ΣΥΝΘΕΤΟΥ ΜΟΝΤΕΛΟΥ ΜΕ ΤΗΝ ΧΡΗΣΗ ΤΟΥ PLG2....	40
ΕΙΚΟΝΑ 7: ΣΥΝΘΕΤΟ ΜΟΝΤΕΛΟ ΒΡΜΝ ΓΙΑ ΤΗΝ ΔΙΕΞΑΓΩΓΗ ΤΩΝ ΠΕΙΡΑΜΑΤΩΝ	40
ΕΙΚΟΝΑ 8: PETRI NET ΑΠΟ ΤΟ ALPHA MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 500 ΕΚΤΕΛΕΣΕΙΣ ΚΑΙ ΧΩΡΙΣ ΘΟΡΥΒΟ	41
ΕΙΚΟΝΑ 9: HEURISTIC NET ΑΠΟ ΤΟ MINE A HEURISTIC NET WITH THE HEURISTIC MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 500 ΕΚΤΕΛΕΣΕΙΣ ΚΑΙ ΧΩΡΙΣ ΘΟΡΥΒΟ	42
ΕΙΚΟΝΑ 10: PROCESS TREE ΑΠΟ ΤΟ MINE A PROCESS TREE WITH ETMD PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 500 ΕΚΤΕΛΕΣΕΙΣ ΚΑΙ ΧΩΡΙΣ ΘΟΡΥΒΟ .	42
ΕΙΚΟΝΑ 11: PETRI NET ΑΠΟ ΤΟ MINE PETRI NET WITH INDUCTIVE MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 500 ΕΚΤΕΛΕΣΕΙΣ ΚΑΙ ΧΩΡΙΣ ΘΟΡΥΒΟ .	43
ΕΙΚΟΝΑ 12: PETRI NET ΑΠΟ ΤΟ ALPHA MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 500 ΕΚΤΕΛΕΣΕΙΣ ΚΑΙ ΘΟΡΥΒΟ ΣΤΑ ACTIVITY NAMES.....	44
ΕΙΚΟΝΑ 13: HEURISTIC NET ΑΠΟ ΤΟ MINE A HEURISTIC NET WITH THE HEURISTIC MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 500 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΘΟΡΥΒΟ ΣΤΑ ACTIVITY NAMES.....	44
ΕΙΚΟΝΑ 14: PROCESS TREE ΑΠΟ ΤΟ MINE A PROCESS TREE WITH ETMD PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 500 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΘΟΡΥΒΟ ΣΤΑ ACTIVITY NAMES.....	45
ΕΙΚΟΝΑ 15: PETRI NET ΑΠΟ ΤΟ MINE PETRI NET WITH INDUCTIVE MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 500 ΕΚΤΕΛΕΣΕΙΣ ΚΑΙ ΘΟΡΥΒΟ ΣΤΑ ACTIVITY NAMES.....	46
ΕΙΚΟΝΑ 16: PETRI NET ΑΠΟ ΤΟ ALPHA MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 500 ΜΕ ΘΟΡΥΒΟ ΣΤΑ CONTROL FLOWS	47
ΕΙΚΟΝΑ 17: HEURISTIC NET ΑΠΟ ΤΟ MINE A HEURISTIC NET WITH THE HEURISTIC MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 500 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΘΟΡΥΒΟ ΣΤΑ CONTROL FLOWS	47
ΕΙΚΟΝΑ 18: PROCESS TREE ΑΠΟ ΤΟ MINE A PROCESS TREE WITH ETMD PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 500 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΘΟΡΥΒΟ ΣΤΑ CONTROL FLOWS	48
ΕΙΚΟΝΑ 19: PETRI NET ΑΠΟ ΤΟ MINE PETRI NET WITH INDUCTIVE MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 500 ΕΚΤΕΛΕΣΕΙΣ ΚΑΙ ΘΟΡΥΒΟ ΣΤΑ CONTROL FLOWS	48
ΕΙΚΟΝΑ 20: PETRI NET ΑΠΟ ΤΟ ALPHA MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 500 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΠΛΗΡΗ ΘΟΡΥΒΟ.....	49
ΕΙΚΟΝΑ 21: HEURISTIC NET ΑΠΟ ΤΟ MINE A HEURISTIC NET WITH THE HEURISTIC MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 500 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΘΟΡΥΒΟ ΣΕ ΟΛΟ ΤΟ ΦΑΣΜΑ ΤΟΥ ΜΟΝΤΕΛΟΥ.....	50
ΕΙΚΟΝΑ 22: PROCESS TREE ΑΠΟ ΤΟ MINE A PROCESS TREE WITH ETMD PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 500 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΠΛΗΡΗ ΘΟΡΥΒΟ	51

EΙΚΟΝΑ 23: PETRI NET ΑΠΟ ΤΟ MINE PETRI NET WITH INDUCTIVE MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 500 ΕΚΤΕΛΕΣΕΙΣ ΚΑΙ ΘΟΡΥΒΟ ΜΕ ΠΛΗΡΗ ΘΟΡΥΒΟ.....	51
EΙΚΟΝΑ 24: PETRI NET ΑΠΟ ΤΟ ALPHA MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1000 ΕΚΤΕΛΕΣΕΙΣ ΚΑΙ ΧΩΡΙΣ ΘΟΡΥΒΟ	52
EΙΚΟΝΑ 25: HEURISTIC NET ΑΠΟ ΤΟ MINE A HEURISTIC NET WITH THE HEURISTIC MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1000 ΕΚΤΕΛΕΣΕΙΣ ΧΩΡΙΣ ΘΟΡΥΒΟ	52
EΙΚΟΝΑ 26: PROCESS TREE ΑΠΟ ΤΟ MINE A PROCESS TREE WITH ETMD PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1000 ΕΠΑΝΑΛΗΨΕΙΣ	53
EΙΚΟΝΑ 27: PETRI NET ΑΠΟ ΤΟ MINE PETRI NET WITH INDUCTIVE MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1000 ΕΚΤΕΛΕΣΕΙΣ ΧΩΡΙΣ ΘΟΡΥΒΟ	53
EΙΚΟΝΑ 28: PETRI NET ΑΠΟ ΤΟ ALPHA MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1000 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΘΟΡΥΒΟ ΣΤΑ ACTIVITY NAMES.....	54
EΙΚΟΝΑ 29: HEURISTIC NET ΑΠΟ ΤΟ MINE HEURISTICS NET WITH HEURISTIC MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1000 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΘΟΡΥΒΟ ΣΤΑ ACTIVITY NAMES.....	55
EΙΚΟΝΑ 30: PROCESS TREE ΑΠΟ ΤΟ MINE A PROCESS TREE WITH ETMD PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1000 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΘΟΡΥΒΟ ΣΤΑ ACTIVITY NAMES.....	55
EΙΚΟΝΑ 31: PETRI NET ΑΠΟ ΤΟ MINE PETRI NET WITH INDUCTIVE MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1000 ΕΚΤΕΛΕΣΕΙΣ ΧΩΡΙΣ ΘΟΡΥΒΟ	56
EΙΚΟΝΑ 32: PETRI NET ΑΠΟ ΤΟ ALPHA MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1000 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΘΟΡΥΒΟ ΣΤΑ CONTROL FLOWS.....	56
EΙΚΟΝΑ 33: HEURISTIC NET ΑΠΟ ΤΟ MINE A HEURISTIC NET WITH THE HEURISTIC MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1000 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΘΟΡΥΒΟ ΣΤΑ CONTROL FLOWS	57
EΙΚΟΝΑ 34: PROCESS TREE ΑΠΟ ΤΟ MINE A PROCESS TREE WITH ETMD PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1000 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΘΟΡΥΒΟ ΣΤΑ CONTROL FLOWS	58
EΙΚΟΝΑ 35: PETRI NET ΑΠΟ ΤΟ MINE A PETRI NET WITH INDUCTIVE MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1000 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΘΟΡΥΒΟ ΣΤΑ CONTROL FLOWS	58
EΙΚΟΝΑ 36: PETRI NET ΑΠΟ ΤΟ ALPHA MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1000 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΠΛΗΡΗ ΘΟΡΥΒΟ.....	59
EΙΚΟΝΑ 37: HEURISTIC NET ΑΠΟ ΤΟ MINE A HEURISTIC NET WITH THE HEURISTIC MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1000 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΠΛΗΡΗ ΘΟΡΥΒΟ	60
EΙΚΟΝΑ 38: ΚΟΜΜΑΤΙ ΑΠΟ ΤΟ HEURISTIC NET ΑΠΟ ΤΟ MINE A HEURISTIC NET WITH THE HEURISTIC MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1000 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΠΛΗΡΗ ΘΟΡΥΒΟ.....	60
EΙΚΟΝΑ 39: PROCESS TREE ΑΠΟ ΤΟ MINE A PROCESS TREE WITH ETMD PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1000 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΠΛΗΡΗ ΘΟΡΥΒΟ	61
EΙΚΟΝΑ 40: PETRI NET ΑΠΟ ΤΟ MINE A PETRI NET WITH INDUCTIVE PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1000 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΠΛΗΡΗ ΘΟΡΥΒΟ	61
EΙΚΟΝΑ 41: PETRI NET ΑΠΟ ΤΟ ALPHA MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1500 ΕΚΤΕΛΕΣΕΙΣ.....	62

EΙΚΟΝΑ 42: HEURISTIC NET ΑΠΟ ΤΟ MINE A HEURISTIC NET WITH THE HEURISTIC MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1500 ΕΚΤΕΛΕΣΕΙΣ ΧΩΡΙΣ ΘΟΡΥΒΟ	63
EΙΚΟΝΑ 43: PROCESS TREE ΑΠΟ ΤΟ MINE A PROCESS TREE WITH ETMD PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1500 ΕΚΤΕΛΕΣΕΙΣ ΚΑΙ ΧΩΡΙΣ ΘΟΡΥΒΟ	63
EΙΚΟΝΑ 44: PETRI NET ΑΠΟ ΤΟ MINE PETRI NET WITH INDUCTIVE MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1500 ΕΚΤΕΛΕΣΕΙΣ ΧΩΡΙΣ ΘΟΡΥΒΟ	64
EΙΚΟΝΑ 45: PETRI NET ΑΠΟ ΤΟ MINE PETRI NET WITH INDUCTIVE MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1500 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΘΟΡΥΒΟ ΣΤΑ ACTIVITY NAMES.....	64
EΙΚΟΝΑ 46: HEURISTIC NET ΑΠΟ ΤΟ MINE HEURISTICS NET WITH HEURISTIC MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1500 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΘΟΡΥΒΟ ΣΤΑ ACTIVITY NAMES.....	65
EΙΚΟΝΑ 47: PROCESS TREE ΑΠΟ ΤΟ MINE A PROCESS TREE WITH ETMD PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1000 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΘΟΡΥΒΟ ΣΤΑ ACTIVITY NAMES.....	65
EΙΚΟΝΑ 48: PETRI NET ΑΠΟ ΤΟ MINE PETRI NET WITH INDUCTIVE MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1000 ΕΚΤΕΛΕΣΕΙΣ ΧΩΡΙΣ ΘΟΡΥΒΟ	66
EΙΚΟΝΑ 49: PETRI NET ΑΠΟ ΤΟ ALPHA MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1000 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΘΟΡΥΒΟ ΣΤΑ CONTROL FLOWS.....	66
EΙΚΟΝΑ 50: HEURISTIC NET ΑΠΟ ΤΟ MINE A HEURISTIC NET WITH THE HEURISTIC MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1000 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΘΟΡΥΒΟ ΣΤΑ CONTROL FLOWS	67
EΙΚΟΝΑ 51: PROCESS TREE ΑΠΟ ΤΟ MINE A PROCESS TREE WITH ETMD PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1500 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΘΟΡΥΒΟ ΣΤΑ CONTROL FLOWS	68
EΙΚΟΝΑ 52: PETRI NET ΑΠΟ ΤΟ MINE A PETRI NET WITH INDUCTIVE MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1500 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΘΟΡΥΒΟ ΣΤΑ CONTROL FLOWS	68
EΙΚΟΝΑ 53: PETRI NET ΑΠΟ ΤΟ ALPHA MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1500 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΠΛΗΡΗ ΘΟΡΥΒΟ.....	69
EΙΚΟΝΑ 54: HEURISTIC NET ΑΠΟ ΤΟ MINE A HEURISTIC NET WITH THE HEURISTIC MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1500 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΠΛΗΡΗ ΘΟΡΥΒΟ	70
EΙΚΟΝΑ 55: PROCESS TREE ΑΠΟ ΤΟ MINE A PROCESS TREE WITH ETMD PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1000 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΠΛΗΡΗ ΘΟΡΥΒΟ	71
EΙΚΟΝΑ 56: PETRI NET ΑΠΟ ΤΟ MINE A PETRI NET WITH INDUCTIVE MINER PLUGIN ΓΙΑ ΤΟ ΠΕΙΡΑΜΑ ΜΕ ΤΙΣ 1500 ΕΚΤΕΛΕΣΕΙΣ ΜΕ ΠΛΗΡΗ ΘΟΡΥΒΟ	72
EΙΚΟΝΑ 57 PROCESS MODEL (A).....	73
EΙΚΟΝΑ 58: PROCESS MODEL (B).....	74
EΙΚΟΝΑ 59: PROCESS MODEL (C).....	74
EΙΚΟΝΑ 60: PROCESS MODEL (D).....	75

Κατάλογος Πινάκων

ΠΙΝΑΚΑΣ 1 ΣΗΜΕΙΟΓΡΑΦΙΑ BUSINESS PROCESSES MODELING NOTATION	20
ΠΙΝΑΚΑΣ 2 ΣΗΜΕΙΟΓΡΑΦΙΑ PETRI NET	23
ΠΙΝΑΚΑΣ 3 EVENT LOG	73
ΠΙΝΑΚΑΣ 4 ΜΕΤΡΙΚΕΣ ΠΟΙΟΤΗΤΑΣ - FITNESS	77
ΠΙΝΑΚΑΣ 5 ΜΕΤΡΙΚΕΣ ΠΟΙΟΤΗΤΑΣ - GENERALIZATION	78
ΠΙΝΑΚΑΣ 6 ΜΕΤΡΙΚΕΣ ΠΟΙΟΤΗΤΑΣ - PRECISION	79

Πρόλογος

Η παρούσα διπλωματική εργασία εκπονήθηκε για το μεταπτυχιακό πρόγραμμα σπουδών Πληροφορικά Συστήματα και Υπηρεσίες με κατεύθυνση τα Μεγάλα Δεδομένα και Αναλυτική κατά τα Ακαδημαϊκά έτη 2018-20, υπό την επίβλεψη του Επίκουρου Καθηγητή κ. Δουλκερίδη Χρήστου, στον οποίο θα ήθελα να εκφράσω τις ευχαριστίες μου για την ανάθεση του θέματος και για τη μετάδοση όλων των απαραίτητων γνώσεων.

Τέλος θα ήθελα να ευχαριστήσω την οικογένειά μου για τη συμπαράστασή και στήριξή τους καθ' όλη τη διάρκεια της διπλωματικής μου εργασίας.

Χρήστος Σταματάκης
Αθήνα, Φεβρουάριος 2020

1 Εισαγωγή

Με το πέρασμα των χρόνων, όλο και περισσότερα δεδομένα καταγράφονται και αποθηκεύονται σε πληροφοριακά συστήματα. Η εξόρυξη διαδικασιών στοχεύει στην ανακάλυψη, παρακολούθηση και βελτίωση των πραγματικών διαδικασιών με την εξαγωγή γνώσεων από τα αρχεία καταγραφής συμβάντων (event logs) που είναι άμεσα διαθέσιμα στα σημερινά πληροφορικά συστήματα. Οι τεχνικές διερεύνησης διεργασιών μπορούν να χρησιμοποιηθούν για την αυτόματη παραγωγή μοντέλων διεργασιών από τα αρχεία καταγραφής συμβάντων. Οι διαφορετικές τεχνικές επιστρέφουν διαφορετικά μοντέλα διεργασιών με κύριες διαφορές στις ιδιότητες και στην ποιότητά τους.

Για την χρήση των πληροφοριακών συστημάτων είναι απαραίτητη η μοντελοποίηση των επιχειρηματικών διαδικασιών. Η Μοντελοποίηση Επιχειρηματικών Διαδικασιών προσφέρει μια κοινή ενιαία γλώσσα και μεθοδολογία για την επικοινωνία των διαδικασιών και των πληροφοριών σχετικά με τις διαδικασίες και τους κανόνες απόφασης, είναι ιδανική για την εκπαίδευση νέων ανθρώπων και τη γρήγορη μεταφορά γνώσης, διότι με μια τεκμηριωμένη διαδικασία, κάθε νέο μέλος της ομάδας μπορεί να εκπαιδευτεί πολύ γρήγορα σε σχέση με το τι πρέπει να κάνει για κάθε κατάσταση που μπορεί να αντιμετωπίσει. Επιπλέον ελαχιστοποιεί τους κίνδυνους απώλειας γνώσης και προσωπικού, βοηθάει στην επικοινωνία μεταξύ των στελεχών των εταιριών και μετατρέπει την εμπειρία μιας ομάδας σε τεκμηριωμένες διαδικασίες.

Υπάρχουν όμως και λόγοι που μειώνουν την χρήση των μοντέλων διεργασιών. Η πλειοψηφία των μοντέλων δεν ανταποκρίνονται πλήρως στη πραγματικότητα των επιχειρήσεων που προσπαθούν να περιγράψουν ενώ παράλληλα τα μοντέλα που έχουν δημιουργηθεί χωρίς την χρήση υπολογιστή προσφέρουν μια ουτοπική εικόνα των διεργασιών. Για τον παραπάνω λόγο γίνεται χρήση των πληροφοριακών συστημάτων που έχουν την δυνατότητα να εκτελούν μοντέλα επιχειρηματικών διεργασιών και να παράγουν τα αρχεία γεγονότων. Τα event logs περιέχουν πληροφορίες σχετικά με την εκτέλεση της κάθε δραστηριότητας και τους τόπους μετάβασης από την μία δραστηριότητα σε μια άλλη. Στην συνέχεια για την μελέτη και την βελτίωση των μοντέλων αντλούνται πληροφορίες από τα αρχεία γεγονότων μέσω του process mining.

1. Process mining

1.1 Ορισμός του Process mining

Για την κάλυψη της ανάγκης των επιχειρήσεων για άντληση περισσότερων πληροφοριών σε σχέση με τις επιχειρηματικές τους διαδικασίες αναπτύχθηκε η τεχνική που ονομάζεται εξόρυξη διεργασιών. Η εξόρυξη διεργασιών βασίζεται στην καταγραφή γεγονότων με σκοπό την εξόρυξη γνώσης. Ο στόχος της συγκεκριμένης τεχνικής είναι μέσω της καταγραφής γεγονότων να παρέχει βελτίωση στα συστήματα πληροφοριών.¹

Επισκόπηση

Η χρήση της εξόρυξης διεργασιών γίνεται όταν δεν υπάρχουν επαρκείς πληροφορίες σχετικά με τον τρόπο απόκτησης διαδικασιών με άλλους τρόπους. Η καταγραφή γεγονότων μπορεί επίσης να χρησιμοποιηθεί για να γίνει σύγκριση με άλλα πρότυπα που έχουν διεξαχθεί στο παρελθόν έτσι ώστε να γίνει έλεγχος αν το τρέχον πρότυπο παρουσιάζει συμμόρφωση σε σχέση με το εκάστοτε κατευθυντήριο μοντέλο.²

Κατηγοριοποίηση

Υπάρχουν τρεις βασικές κατηγορίες ταξινόμησης για τις τεχνικές εξόρυξης διαδικασιών. Η κατηγοριοποίηση χρησιμοποιεί ως βάση εάν προϋπήρχε ένα προγενέστερο μοντέλο και, αν ναι, πως αυτό έχει χρησιμοποιηθεί.

- **Ανακάλυψη:** Δεν προϋπάρχει ένα προγενέστερο μοντέλο. Το μοντέλο θα δημιουργηθεί από την αρχή.
- **Συμμόρφωση:** Υπάρχει ένα μοντέλο εκ των προτέρων. Το προϋπάρχον μοντέλο θα συσχετισθεί με ένα αρχείο καταγραφής συμβάντων με την λογική να υπάρξει συμμόρφωση σε τυχόν στρεβλώσεις του μοντέλου.
- **Επέκταση:** Όπως και στην συμμόρφωση υπάρχει ένα μοντέλο εκ των προτέρων. Το μοντέλο αυτό εξελίσσεται και ο στόχος, σε αντίθεση με τη συμμόρφωση, είναι να βελτιωθεί το μοντέλο.³

¹ Wil Van Der Aalst (2016) Process Mining: Data Science in Action 2nd edition Springer

² Wil Van Der Aalst (2016) Process Mining: Data Science in Action 2nd edition Springer

³ el.wikipedia.org/wiki/Εξόρυξη_διεργασιών

Επιχειρηματικό Μοντέλο

Ο όρος επιχειρηματικό μοντέλο χρησιμοποιείται για την περιγραφή διαφόρων τομέων που σχετίζονται με βασικές έννοιες των επιχειρηματικών δραστηριοτήτων όπως την αποστολή, την προσφορά ή ακόμη και δραστηριότητες που σχετίζονται με την στρατηγική οργάνωσης εμπορικών, πολιτικών και άλλων πρακτικών. Ένα επιχειρηματικό μοντέλο περιγράφει τον τρόπο με τον οποίο ασκείται μία επιχειρηματική δραστηριότητα και τον τρόπο με τον οποίο παράγει έσοδα.

Επιχειρησιακή Διεργασία

Επιχειρηματική διεργασία ονομάζεται μία σειρά από διεργασίες που σχετίζονται μεταξύ τους και παράγουν κάποιο συγκεκριμένο προϊόν ή κάποια υπηρεσία.

Μια επιχειρησιακή διεργασία μπορεί να διασπαστεί σε μικρότερες διεργασίες με δικιές τους ιδιότητες οι οποίες όμως λειτουργούν για την διεκπεραίωση του στόχου της αρχικής διεργασίας.

Το μοντέλο επιχειρησιακής διεργασίας είναι ένα μοντέλο μιας ή περισσότερων επιχειρησιακών διεργασιών και καθορίζει τους τρόπους με τους οποίους πραγματοποιούνται οι διάφορες λειτουργίες για να εκπληρωθούν οι επιδιωκόμενοι στόχοι ενός οργανισμού. Ένα τέτοιο μοντέλο παραμένει μία αφηρημένη έννοια και εξαρτάται από την προβλεπόμενη χρήση του.⁴

Αλγόριθμοι που χρησιμοποιούνται για Process Mining

Το Process Mining έχει να κάνει με την χρήση διαφόρων αλγορίθμων για την ανακάλυψη μοντέλων, κυρίως με τους εξής: Alpha Miner, Heuristic Miner, Evolutionary Tree Miner (Genetic Miner) και Inductive Miner. Οι αλγόριθμοι αυτοί είναι αλγόριθμοι που έχουν την δυνατότητα ανακάλυψης διεργασιών καθώς μπορούν να δημιουργήσουν από την αρχή μοντέλα διεργασιών μέσω των event log.

⁴ Burkhart, Thomas; Krumeich, Julian; Werth, Dirk; and Loos, Peter, "Analyzing the Business Model Concept — A Comprehensive Classification of Literature" (2011). ICIS 2011 Proceedings.

1.2 Εφαρμογές του Process Mining

Τα πληροφορικά συστήματα γίνονται διαρκώς όλο και πιο αλληλένδετα μεταξύ τους, με βάση τις παραγωγικές δραστηριότητες τις οποίες υποστηρίζουν. Όπως είναι γνωστό τεράστιος αριθμός γεγονότων καταγράφεται σε διάρκεια μιας μέρας από τα πληροφορικά συστήματα. Παρόλα αυτά οι επιχειρήσεις έχουν δυσκολία να εξάγουν την προστιθέμενη αυτή αξία από τα δεδομένα τα οποία δημιουργούνται. Ο κύριος στόχος του process mining είναι να χρησιμοποιήσει αυτά τα καταγεγραμμένα γεγονότα δημιουργώντας μοντέλα διεργασιών. Η μοντελοποίηση αυτή χρησιμοποιείται από επιχειρησιακούς αναλυτές τους μάνατζερ και όσους εργάζονται πάνω στην βελτιστοποίηση των διεργασιών. Με το πέρασμα του χρόνου τα πλεονεκτήματα από την εξέλιξη της τεχνολογίας δίνουν την δυνατότητα καταγραφής μοντέλων Επιχειρησιακής Μοντελοποίησης που έχουν την ικανότητα να προσομοιωθούν.

Μοντελοποίηση και Προσομοίωση

Δεν είναι εύκολο να δημιουργήσεις ένα καλό επιχειρησιακό μοντέλο, όμως είναι σημαντικό. Το process mining είναι η λύση καθώς μπορεί να κατασκευάσει καλύτερα μοντέλα σε λιγότερο χρόνο. Οι αλγόριθμοι εξερεύνησης διεργασιών όπως ο α -algorithm μπορεί να αυτοματοποιήσει την μοντελοποίηση των διεργασιών. Ο σκοπός της μοντελοποίησης είναι να αποφασίσει ποιες δραστηριότητες χρειάζεται να εκτελεστούν και με ποια σειρά κατά την διάρκεια της προσομοίωσης.

Ορισμένες τεχνικές επιχειρηματικής μοντελοποίησης είναι οι ακόλουθες:

- Transition Systems
- Petri Nets
- Workflow Nets
- YAWL (Yet Another Workflow Language)
- Business Process Modeling Notation (BPMN)
- Event-Driven Process Chains (EPCs)
- Casual Nets
- Process Trees⁵

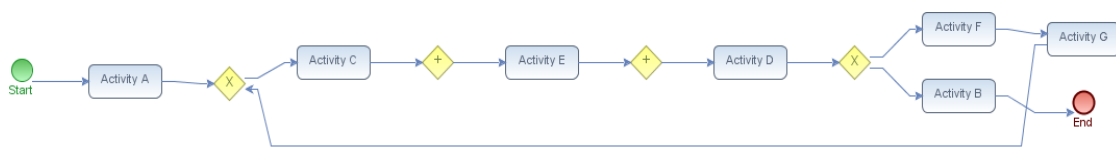
⁵ Wil Van Der Aalst (2016) Process Mining: Data Science in Action 2nd edition Springer

1.3 Μοντέλα Διεργασιών και Δεδομένα

Η πληθώρα σημειογραφία των μοντέλων διεργασιών δείχνει τη σημασία της μοντελοποίησης των διαδικασιών. Στην παρούσα εργασία θα παρουσιαστούν τα μοντέλα διεργασιών που χρησιμοποιήθηκαν.

Business Process Modeling Notation – BPMN

Ο οργανισμός BPMI ανέπτυξε το BPMN το οποίο είναι ένα κοινό πρότυπο αναπαράστασης των επιχειρηματικών διαδικασιών και επιτρέπει την ανταλλαγή σχετικών πληροφοριών ανάμεσα σε επιχειρήσεις με επιχειρήσεις αλλά και ανάμεσα σε επιχειρήσεις με πελάτες. Πρόσφατα το BPMN η αλλιώς το Business Process Modeling Notation έχει γίνει η πιο ευρέως χρησιμοποιημένη γλώσσα μοντέλων επιχειρησιακών διεργασιών.⁶



Εικόνα 1 Μοντέλο με σημειογραφία του BPMN

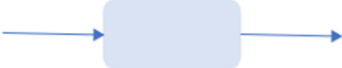



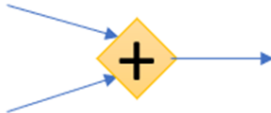
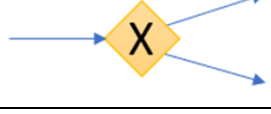
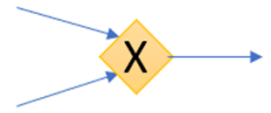
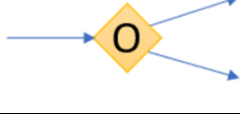
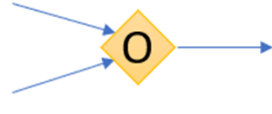
Στην εικόνα 1, παρατηρείται ένα μικρό υποσύνολο από στοιχεία της σημειογραφίας. Η μεγαλύτερη διαφορά σε σχέση με άλλες τεχνικές επιχειρηματικής μοντελοποίησης είναι ότι η λογική δρομολόγησης δηλαδή η ροή δεν σχετίζεται με τις δραστηριότητες (Tasks) αλλά με τις διαφορετικές πύλες (Gateways). Υπάρχουν πολλών ειδών διαφορετικές πύλες που είτε διαχωρίζουν είτε ενώνουν διαφορετικές δραστηριότητες (AND, XOR, OR). Τα χωρίσματα αυτά είναι βασισμένα πάνω στις συνθήκες των δεδομένων. Ένα γεγονός (event) είναι συγκρίσιμο με μια περιοχή – δραστηριότητα πάνω σε ένα Petri Net κάτι που δεν ισχύει για το BPMN. Δεν χρειάζεται να εισάγονται γεγονότα (event) ενδιάμεσα στις δραστηριότητες όπως επίσης τα γεγονότα δεν μπορούν να έχουν πολλαπλές εισαγωγές και εξαγωγές. Τα αρχικά γεγονότα έχουν μια εξαγωγή, τα ενδιάμεσα μια εξαγωγή και μία εισαγωγή και το τελικό γεγονός μια εισαγωγή. Όπως προαναφέρθηκε τα γεγονότα δεν μπορούν να

⁶ Stephen A. White, (2004) "Business Process Management Initiative (BPMI) Business Process Modeling Notation (BPMN)" Copyright 2004, BPMI.org.

έχουν περισσότερες από μία εισαγωγές για τον λόγο αυτό οι σύνδεσμοι και τα χωρίσματα γίνονται χρησιμοποιώντας τις πύλες.

Σημειογραφία στο πρότυπο της BPMN

Παρακάτω θα παρουσιαστεί η σημειογραφία του μοντέλου BPMN που χρησιμοποιήσαμε.

Σημειογραφία	
	Task/Activity
	Start Event
	End Event
	AND-split gateway
	AND-join gateway
	XOR-split gateway
	XOR-join gateway
	OR-split gateway
	OR-join gateway

Πίνακας 1 Σημειογραφία Business Processes Modeling Notation⁷

⁷ Mark von Rosing, August-Wilhelm Scheer and Henrik von Scheel (2015) "Business Process Model and Notation—BPMN"

Ορισμός Event Log

Το event log είναι ένα αρχείο καταγραφής συμβάντων(events) όπου κάθε συμβάν αναφέρεται σε μια περίπτωση, με κάποια δραστηριότητα σε κάποιο χρονικό σημείο. Στην συγκεκριμένη διπλωματική εργασία γίνεται εξόρυξη διεργασιών με την εφαρμογή διαφόρων σχετικών αλγορίθμων σε αρχεία συμβάντων (event logs) που παράγονται από την εκτέλεση κάποιων μοντέλων BPMN. Τα παραγόμενα event log είναι σε μορφή XES (eXtensible Event Stream) και εμπεριέχουν πεδία σχετικά με την κάθε εκτέλεση (Trace) των BPMN. Για το κάθε trace υπάρχει αναγνωριστικό πεδίο με το κλειδί concept:name.

```
<extension name="Time" prefix="time" uri="http://www.xes-standard.org/time.xesext"/>
<extension name="Lifecycle" prefix="lifecycle" uri="http://www.xes-standard.org/lifecycle.xesext"/>
<extension name="Concept" prefix="concept" uri="http://www.xes-standard.org/concept.xesext"/>
<string key="concept:name" value="tmp-process"/>
<trace>
  <string key="concept:name" value="case_971"/>
  <event>
    <string key="concept:name" value="Activity A"/>
    <date key="time:timestamp" value="1970-01-01T02:00:00+02:00"/>
  </event>
  <event>
    <string key="concept:name" value="Activity D"/>
    <string key="variable_b" value="test-value-16"/>
    <date key="time:timestamp" value="1970-01-01T03:00:00+02:00"/>
  </event>
  <event>
    <string key="concept:name" value="Activity B"/>
    <date key="time:timestamp" value="1970-01-01T04:00:00+02:00"/>
  </event>
</trace>
<trace>
  <string key="concept:name" value="case_333"/>
  <event>
    <string key="concept:name" value="Activity A"/>
    <date key="time:timestamp" value="1970-01-01T02:00:00+02:00"/>
  </event>
  <event>
    <string key="concept:name" value="Activity C"/>
    <string key="variable_a" value="fkt9oli5nf48h"/>
    <date key="time:timestamp" value="1970-01-01T03:00:00+02:00"/>
  </event>
  <event>

```

Εικόνα 2 Traces Event Log από εκτέλεση BPMN (Πεδίο Trace).

Μέσα σε κάθε trace εμπεριέχεται πληροφορία σχετικά με το κάθε event που εκτελέστηκε για το οποίο επίσης υπάρχει αναγνωριστικό πεδίο concept:name και η αντίστοιχη ημερομηνία και ώρα στο πεδίο time:timestamp.

```

<string key="concept:name" value="tmp-process"/>
<trace>
  <string key="concept:name" value="case_971"/>
  <event>
    <string key="concept:name" value="Activity A"/>
    <date key="time:timestamp" value="1970-01-01T02:00:00+02:00"/>
  </event>
  <event>
    <string key="concept:name" value="Activity D"/>
    <string key="variable_b" value="test-value-16"/>
    <date key="time:timestamp" value="1970-01-01T03:00:00+02:00"/>
  </event>
  <event>
    <string key="concept:name" value="Activity B"/>
    <date key="time:timestamp" value="1970-01-01T04:00:00+02:00"/>
  </event>
</trace>
<trace>
  <string key="concept:name" value="case_333"/>
  <event>
    <string key="concept:name" value="Activity A"/>
    <date key="time:timestamp" value="1970-01-01T02:00:00+02:00"/>
  </event>
  <event>
    <string key="concept:name" value="Activity C"/>
    <string key="variable_a" value="fkt9o1i5nf48h"/>
    <date key="time:timestamp" value="1970-01-01T03:00:00+02:00"/>
  </event>
  <event>
    <string key="concept:name" value="Activity B"/>
    <date key="time:timestamp" value="1970-01-01T04:00:00+02:00"/>
  </event>
</trace>

```

Εικόνα 3 Traces Event Log από εκτέλεση BPMN (Πεδίο timestamp)

Petri Nets

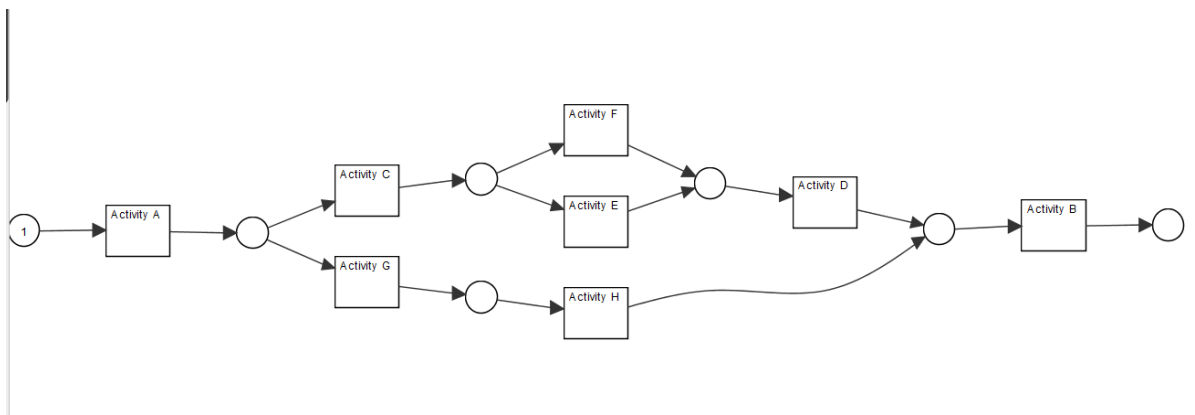
Το Petri Net είναι ένα εργαλείο το οποίο μέσω μαθηματικών θεμελίων έχει την δυνατότητα να αναπαραστήσει αλληλεπιδράσεις ενεργών και παθητικών συστατικών. Τα οποία είναι οι θέσεις (ενεργά) και οι μεταβάσεις (παθητικά). Οι σχέσεις μεταξύ των συστατικών συμβολίζονται με τόξα. Για την γραφική αναπαράσταση της πληροφορίας μέσα στα PN χρησιμοποιούνται τα γραφικά αντικείμενα που ονομάζονται κουπόνια. Για την ακρίβεια είναι σημεία που μετακινούνται μεταξύ των θέσεων μέσω των μεταβάσεων. Για την αναπαράσταση των μεταβάσεων χρησιμοποιούνται ορθογώνια και των θέσεων κύκλοι. Τα Petri Nets είναι ιδανικά για την αναπαράσταση συστημάτων όπου οι εξαρτήσεις μπορεί να αλλάζουν δυναμικά και για περιπτώσεις συστημάτων με ταυτόχρονη επεξεργασία.

Χαρακτηριστικά των Petri Nets




- Δίνουν την δυνατότητα γραφικής αναπαράστασης συστημάτων όπου υπάρχουν εξαρτήσεις

- Δίνουν την δυνατότητα να περιγραφεί ένα σύστημα μέσω βαθμών αφαίρεσης
- Δίνουν την δυνατότητα γραφικής αναπαράστασης με μέσα όμοια του σύστημα που αναπαριστούν

Τα δίκτυα PetriNets είναι η παλαιότερη και πιο διαδεδομένη γλώσσα μοντελοποίησης της διαδικασίας διερεύνησης, η οποία επιτρέπει τη μοντελοποίηση των συχνοτήτων. Παρόλο που η γραφική σημειογραφία είναι εξαιρετικά απλή, τα PN είναι εκτελέσιμα και έχουν μεγάλες δυνατότητες αναλυτικών τεχνικών για την καλύτερη ανάλυση τους.⁸



Εικόνα 4 Petri Net

Σημειογραφία	
	Transition
	Place
	Token

Πίνακας 2 Σημειογραφία Petri Net

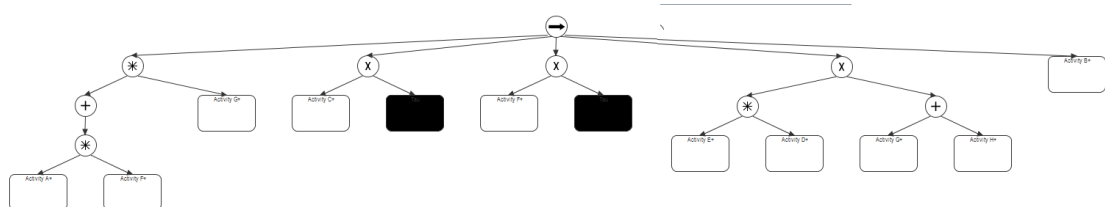
Τα Petri Nets είναι ένα διμερές γράφημα που αποτελείται από θέσεις και μεταβάσεις. Η δομή του δικτύου είναι στατική, αλλά, σύμφωνα με τον κανόνα πυροδότησης, τα tokens η αλλιώς οι εκτελέσεις μπορούν να ρέουν μέσω του δικτύου.

⁸ Cardoso, Janette, Camargo, Heloisa, (1999) " Fuzziness in Petri Nets " Springer

Τα PNs όπως αναφέρθηκε έχουν ένα πολύ ισχυρό θεωρητικό υπόβαθρο που μπορεί να καλύψει την πληροφορία του συγχρονισμού και των συχνοτήτων πολύ καλά. Είναι προφανές ότι αυτά μοντέλα είναι περιεκτικά σε πληροφορία και έχουν μια αδυναμία να εκφράσουν προβλήματα που έχουν να κάνουν με το μέγεθος των δεδομένων όπως επίσης παρουσιάζουν αδυναμία να εκφράσουν τις πληροφορίες που έχουν σχέση με τον χρόνο. Για τον λόγο αυτό έχουν προταθεί Petri Nets μεγαλύτερου level δηλαδή πιο εξειδικευμένα. Τα πιο διαδεδομένα Petri Nets είναι τα χρωματιστά Petri Nets. Τα Colored Petri Nets (CPNs) είναι τα πιο διαδεδομένα καθώς μπορούν να εξιδανικεύσουν την πληροφορία σε σχέση με τον όγκο των δεδομένων δηλαδή πόσες φορές χρησιμοποιήθηκε μία δραστηριότητα και την χρονική διάρκεια που είχε η κάθε δραστηριότητα.⁹

Process Trees

Σε αντίθεση με άλλα μοντέλα διεργασίας όπως το BPMN, το Petri Net και το YAWL, το Process Tree δεν έχει αδυναμίες σε αδιέξοδα, loops ή παράλληλες ενέργειες. Τα προαναφερθέντα μοντέλα που έχουν μη επιθυμητές ιδιότητες ανεξαρτήτως των event logs ονομάζονται επισφαλής (unsound). Έτσι ορισμένες φορές δεν χρειάζεται να ανατρέξουμε στο αρχείο καταγραφής των συμβάντων (event log) για να διαπιστώσουμε ότι το μοντέλο μας δεν μπορεί να περιγράψει σωστά την παρατηρούμενη συμπεριφορά. Αυτές οι ανωμαλίες δεν εμφανίζονται στο Process Tree. Τα δέντρα διεργασιών μπορούν να θεωρηθούν ως ιεραρχικά μοντέλα διεργασιών στα οποία οι κόμβοι (Nodes) είναι οι χειριστές της ακολουθίας και της επιλογής των δραστηριοτήτων και τα φύλλα είναι η δραστηριότητες.



Εικόνα 5 Process Tree

⁹ Wil Van Der Aalst (2016) Process Mining: Data Science in Action 2nd edition Springer

Ένα δέντρο διαδικασίας είναι ένα σχηματικό διάγραμμα των διαδικασιών που ένα προϊόν περνάει κατά τη διάρκεια της ζωής του. Μεταξύ της δημιουργίας και της διάθεσής του, ένα προϊόν περνά μέσα από διαδικασίες όπως η κατασκευή, συναρμολόγηση, διανομή, εγκατάσταση, λειτουργία, συντήρηση, χρήση, επαναχρησιμοποίηση και διάθεση. Κάθε μία από αυτές τις διαδικασίες έρχεται με ορισμένες απαιτήσεις για το νέο μοντέλο. Δημιουργώντας ένα δέντρο διαδικασιών αναγκάζεται ο σχεδιαστής να σκεφτεί μελλοντικά: σε ποιες καταστάσεις, τόπους, δραστηριότητες θα εμφανιστεί το νέο προϊόν, ποιος κάνει τι με το προϊόν, ποια προβλήματα αναμένονται, ποιες απαιτήσεις απαιτούνται για αυτές τις καταστάσεις. Ένα δέντρο διαδικασιών αναγκάζει τον σχεδιαστή να σκέφτεται συστηματικά μέσω όλων των υπό-έργων που περνά ένα προϊόν: παραγωγή (συμπεριλαμβανομένης της ανάπτυξης), διανομή, χρήση και διάθεση.¹⁰

¹⁰ Wil Van Der Aalst (2016) Process Mining: Data Science in Action 2nd edition Springer

2 Περιγραφή Process Mining Software

2.1 Λογισμικό Δημιουργίας Τυχαίων Διεργασιών PLG2

Η αξιολόγηση των αλγορίθμων εξόρυξης διεργασιών απαιτεί την διαθεσιμότητα μιας σειράς επιχειρησιακών διαδικασιών πραγματικού κόσμου και των αρχείων καταγραφής εκτέλεσης τους, τα οποία δύσκολα είναι διαθέσιμα.

Ένας από τους σημαντικότερους λόγους που επιλέχθηκε να χρησιμοποιηθεί το PLG2 στην τρέχουσα διπλωματική εργασία είναι γιατί πρόκειται για μια εφαρμογή ικανή να παράγει τυχαίες επιχειρηματικές διαδικασίες ξεκινώντας από κάποιες γενικές "παραμέτρους πολυπλοκότητας". Το PLG2 είναι επίσης σε θέση να "εκτελέσει" ένα μοντέλο διεργασίας για να δημιουργήσει ένα αρχείο καταγραφής.

Αυτό το λογισμικό έχει σχεδιαστεί για να βοηθήσει τους ερευνητές στην κατασκευή ενός μεγάλου συνόλου διαδικασιών και αντίστοιχων αρχείων καταγραφής εκτέλεσης. Αυτό το λογισμικό κυκλοφορεί με μια μικρή βιβλιοθήκη που θα μπορούσε να βοηθήσει στην προγραμματική δημιουργία διαδικασιών.

Χαρακτηριστικά του PLG2

- Τυχαία παραγωγή και προσομοίωση σύνθετων μοντέλων διεργασιών
- Τυχαία παραγωγή και προσομοίωση σύνθετων αντικειμένων δεδομένων
- Λεπτομερής αλλαγές στα χαρακτηριστικά του χρόνου δραστηριοτήτων
- Τυχαία εξέλιξη ενός μοντέλου διεργασίας, προκειμένου να δημιουργηθεί μια ελαφρώς διαφορετική εκδοχή του
- Έλεγχος του θορύβου προσομοίωσης
- Εισαγωγή μοντέλου διαδικασίας από αρχεία XML PLG2 ή BPMN
- Εξαγωγή μοντέλων σε μορφές BPMN 2.0 XML, PNML και PLG2¹¹

¹¹ Andrea Burattin (2016) "PLG2: Multiperspective Process Randomization with Online and Offline Simulations"

2.2 Το εργαλείο ProM

Είναι αδύνατο να δοθεί μία ολοκληρωμένη εικόνα για όλα τα προγράμματα που εξυπηρετούν την εξόρυξη διεργασιών. Το ProM είναι ένα εργαλείο λογισμικού που δίνει την δυνατότητα ανάλυσης και παρακολούθησης πραγματικών διαδικασιών. Είναι ένα επεκτάσιμο framework ανοιχτού κώδικα που υποστηρίζει πολλές process mining τεχνικές μέσω των plugin του. Το ProM ήδη μετράει περισσότερες από 1500 εισαγωγές (Plug-ins). Αυτά τα plug-ins εξυπηρετούν ένα τεράστιο εύρος από τεχνικές ανάλυσης. Για παράδειγμα όλες οι τεχνικές εξόρυξης διεργασιών στην συγκεκριμένη εργασία υποστηρίχθηκαν μέσα από τα plug-ins του ProM. Τα όρια του ProM δεν σταματάνε μόνο στα όρια της διερεύνησης διεργασιών αλλά επεκτείνονται στον έλεγχο συμμόρφωσης, στην ανάλυση κοινωνικών δικτύων, την ανάλυση της συμμόρφωσης, εξόρυξη αποφάσεων, λειτουργική υποστήριξη, την επαλήθευση και τέλος την μετατροπή του μοντέλου. Σε σχέση με άλλες εφαρμογές το ProM δεν παρουσιάζει αδυναμίες να εντοπιστεί καλά τη συνοχή ενός μοντέλου και παρουσιάζει περιορισμένη υποστήριξη του ελέγχου συμμόρφωσης. Το ProM δεν επικεντρώνεται μόνο στην επίδοση της ανάλυσης αλλά και στον έλεγχο της συμμόρφωσης όπως και στην ακρίβεια των μοντέλων.¹²

Χρήση του Prom

Το ProM μπορεί να χρησιμοποιηθεί για την εύρεση των βημάτων μια διεργασίας, την εύρεση των εξαρτήσεων μεταξύ των διεργασιών, των τρόπων εκτέλεσης τους (παράλληλα ή όχι), την προτεραιότητα τους αλλά και την εύρεση των κεντρικών ή συντονιστικών διεργασιών. Επίσης με την χρήση του ProM μπορούν να παραχθεί πληροφορία σε σχέση με μοτίβα εκτελέσεων και συχνότητων.

Τα αρχεία που χρησιμοποιούνται ως είσοδος για την Χρήση του ProM είναι τα αρχεία τύπου Xes. Ένα event log είναι ένα αρχείο καταγραφής συμβάντων (events) όπου κάθε συμβάν αναφέρεται σε μια περίπτωση, με κάποια δραστηριότητα σε κάποιο χρονικό σημείο. Τα event log είναι σε μορφή XES (eXtensible Event Stream) και εμπεριέχουν πεδία σχετικά με την κάθε εκτέλεση

¹² processmining.org/prom

(Trace) των BPMN. Για το κάθε trace υπάρχει αναγνωριστικό πεδίο με το κλειδί concept:name.

Στην παρούσα διπλωματική εργασία θα γίνει χρήση των Prom Plugins για την εκτέλεση των αλγορίθμων Alpha Miner , Evolutionary Tree Miner, Heuristic Miner και Inductive Miner και θα διεξαχθούν οι μετρικές ποιότητας Fitness, Precision και Generalization για την σύγκριση και μελέτη της ποιότητας των μοντέλων.¹³

¹³ Wil Van Der Aalst (2016) Process Mining: Data Science in Action 2nd edition Springer

3 Αλγόριθμοι και μετρικές ποιότητας

3.1 Αλγόριθμοι

Η ανακάλυψη διεργασιών είναι ένα από τα πιο δύσκολα έργα στην εξόρυξη διεργασιών. Βασισμένο σε ένα event log ένα μοντέλο διεργασιών είναι κατασκευασμένο να αποτυπώνει την διαφαινόμενη συμπεριφορά του log αρχείου. Ο συνδυασμός της εξερεύνησης και της ροής ελέγχου είναι τα κύρια συστατικά της ανακάλυψης διεργασιών. Η διατύπωση του προβλήματος της ανακάλυψης διεργασιών μπορεί να αναφερθεί ως εξής:

Ορίζουμε L ένα event log αρχείο με την μορφή XES όπως ορίστηκε στο κεφάλαιο 1. Ένας αλγόριθμος ανακάλυψης διεργασιών είναι η εξίσωση που χαρτογραφεί το αρχείο L σε ένα μοντέλο διεργασιών έτσι ώστε το μοντέλο να είναι αντιπροσωπευτικό για τη συμπεριφορά που εμφανίζεται στο αρχείο καταγραφής συμβάντων. Η πρόκληση είναι να βρούμε έναν τέτοιο αλγόριθμο. Στο κεφάλαιο αυτό παρουσιάζονται οι αλγόριθμοι που χρησιμοποιήθηκαν στην παρούσα εργασία.¹⁴

3.1.1 Alpha Miner

Αυτός ήταν ο πρώτος αλγόριθμος ανακάλυψης και χρησίμευσε ως βάση για την ανάπτυξη μεταγενέστερων βελτιωμένων αλγορίθμων. Ο βασικός περιορισμός του είναι ότι δεν χρησιμοποιεί συχνότητες και επομένως δεν εγγυάται την ευρωστία οπότε είναι κατάλληλος μόνο για event logs χωρίς θόρυβο το οποίο είναι πολύ σπάνιο για αντίστοιχες διαδικασίες. Ο αλγόριθμος alpha miner κληρονομεί μέσω του event log ένα petri net. Οι κανόνες για την κληρονομήση του petri net είναι οι εξής:

Προσωρινή εξάρτιση (a->b) : Το b ακολουθεί το a και το a δεν ακολουθεί ποτέ το b. Το b είναι εξαρτώμενο από το a.

Παράδειγμα από event log : [(a,b,c,d)(e,d,a,b)(a,e)]

¹⁴ Wil Van Der Aalst (2016) Process Mining: Data Science in Action 2nd edition Springer

Προσωρινή ανεξαρτησία (a||b): Το a σε κάποια διαδρομή έχει ακολουθήσει το b και το b σε κάποια διαδρομή έχει ακολουθήσει a. Το a και το b μπορούν να εκτελεστούν ταυτόχρονα.

Παράδειγμα από event log: [(a,b,c,d)(e,d,b,a)(a,e)]

Ανεξαρτησία (a#b): Το a και το b είναι τελείως ανεξάρτητα

Παράδειγμα από event log : [(a,c,d,b)(a,e,d,b)(a,e)]

Ο μαθηματικός τύπος του αλγόριθμου Alpha Miner

Για να επισημοποιήσουμε τυπικά τον α-αλγόριθμο πρέπει να εισάγουμε κάποια βασική ορολογία. Ορίζουμε T ένα σύνολο δραστηριοτήτων και $\sigma = a_1 a_2 \dots a_n \in T^*$ την αλληλουχία του T με μέγεθος n με το πρώτο και τελευταίο να ορίζεται ως εξής:

1. $a \in \sigma$ αν και μόνο αν $a \in \{a_1, a_2, \dots, a_n\}$,
2. Αν $n \geq 1$, τότε το πρώτο $(\sigma) = a_1$ και το τελευταίο $(\sigma) = a_n$.

Τώρα έχουμε την δυνατότητα να ορίσουμε τον α-αλγόριθμο με μια πιο διαισθητική εξήγηση.

Ορίζουμε W είναι ένα αρχείο workflow πάνω από T . Ο αλγόριθμος $\alpha(W)$ ορίζεται ως εξής.

1. $TW = \{t \in T \mid \exists \sigma \in W t \in \sigma\}$,
2. $TI = \{t \in T \mid \exists \sigma \in W t = \text{first}(\sigma)\}$,
3. $TO = \{t \in T \mid \exists \sigma \in W t = \text{last}(\sigma)\}$,
4. $XW = \{(A,B) \mid A \subseteq TW \wedge B \subseteq TW \wedge \forall a \in A \forall b \in B a \rightarrow W b \wedge \forall a_1, a_2 \in A a_1 \# W a_2 \wedge \forall b_1, b_2 \in B b_1 \# W b_2\}$,
5. $YW = \{(A,B) \in XW \mid \forall (A_-, B_-) \in XW A \subseteq A_- \wedge B \subseteq B_- \implies (A,B) = (A_-, B_-)\}$,
6. $PW = \{p(A,B) \mid (A,B) \in YW\} \cup \{iW, oW\}$,
7. $FW = \{(a, p(A,B)) \mid (A,B) \in YW \wedge a \in A\} \cup \{(p(A,B), b) \mid (A,B) \in YW \wedge b \in B\} \cup \{(iW, t) \mid t \in TI\} \cup \{(t, oW) \mid t \in TO\}$, and
8. $\alpha(W) = (PW, TW, FW)$.¹⁵

¹⁵ Ana Karla A. de Medeiros, Wil Van der Aalst, A. J. M. M. Weijters, (2003), "Workflow Mining: Current Status and Future Directions"

Περιγραφή των βημάτων

1. Ορισμός όλων των events
2. Ορισμός όλων των πιθανών αρχικών events
3. Ορισμός όλων των πιθανών τελικών events
4. Υπολογισμός των πιθανών συνόλων A και B. Όλα τα events που βρίσκονται μέσα στο σύνολο A και όλα τα events που βρίσκονται στο σύνολο B πρέπει να είναι ανεξάρτητα μεταξύ τους (Τα events του ίδιου συνόλου να είναι ανεξάρτητα μεταξύ τους). Όλα τα events του A συνόλου πρέπει να είναι προσωρινά συσχετισμένα με τα events του συνόλου B.
5. Απαλειφή των μη μέγιστων συνόλων
6. Δημιουργία θέσεων για όλα τα κληρονομημένα ορίσματα
7. Σχεδιασμός των συνδέσεων
8. Επιστροφή του Petri Net¹⁶

3.1.2 Heuristic Miner

Οι ευριστικοί αλγόριθμοι εξόρυξης χρησιμοποιούν μια αναπαράσταση παρόμοια με την Casual Nets. Επιπλέον, αυτοί οι αλγόριθμοι λαμβάνουν υπόψη τις συχνότητες γεγονότων και αλληλουχιών κατά την κατασκευή του μοντέλου. Η ιδέα είναι ότι οι σπάνιες διαδρομές δεν θα πρέπει να ενσωματωθούν στο μοντέλο. Τόσο η μοντελοποίηση μέσω των Casual Nets όσο και η χρήση των συχνότητων κάνει την προσέγγιση τους να μοιάζει πιο ισχυρή από τις υπόλοιπες προσεγγίσεις. Στα πειράματα που έλαβαν μέρος στην παρούσα διπλωματική εργασία τα Casual Nets η αλλιώς Heuristic Nets μετατράπηκαν σε Petri Nets με την βοήθεια των Plugins του ProM.

Heuristic Miner (HM): Έχει τρεις σημαντικές βελτιώσεις σε σχέση με τον αλγόριθμο Alpha. Πρώτον, λαμβάνει υπόψη τις συχνότητες και την σημασία, ώστε να μπορεί να καταργήσει θορυβώδη ή σπάνια συμπεριφορά, γεγονός που καθιστά λιγότερο ευαίσθητο στον θόρυβο. Δεύτερον, μπορεί να ανιχνεύσει βραχυκύκλωμα. Τρίτον, επιτρέπει να παραλειφθούν οι μεμονωμένες δραστηριότητες.

Τα τρία βήματα για την εξόρυξη διεργασιών του Heuristic Miner:

1. Εμφάνιση εξαρτήσεων των συγγενών δραστηριοτήτων.

¹⁶ Ana Karla A. de Medeiros, Wil Van der Aalst, A. J. M. M. Weijters, (2003), "Workflow Mining: Current Status and Future Directions"

Στο βήμα αυτό ο αλγόριθμος Heuristic Miner προσπαθεί να ποσοτικοποιήσει την εξάρτηση ανάμεσα στις δραστηριότητες θέτοντας μια συνάρτηση βάρους w ανάμεσα τους. Εστω δραστηριότητα A και B τότε η εξάρτηση μεταξύ του διαφαίνεται ως $A \rightarrow wB$. Όσο μεγαλύτερη είναι η εξάρτηση ανάμεσα σε δύο δραστηριότητες ο αλγόριθμος τις θεωρεί σαν κύρια ροή και την ενσωματώνει στο Heuristic Net που δημιουργεί.

2. Καταγραφή παράλληλων δραστηριοτήτων.

Δύο δραστηριότητες οι οποίες εμφανίζονται μέσα στο αρχείο καταγραφής (event log) σαν αλληλουχία ...ΓΔ... τότε μπορεί να έχουν την σχέση AND. Από την άλλη μεριά δεν μπορεί να υπάρξει αλληλουχία δραστηριοτήτων ...ΓΔ... μέσα στο αρχείο και οι δραστηριότητες αυτές να έχουν σχέση OR ή XOR.

3. Εξάρτηση δραστηριοτήτων μεγάλων αποστάσεων.

Στο τρίτο και τελευταίο βήμα ο αλγόριθμος προσπαθεί να καταγράψει τις εξαρτήσεις ανάμεσα σε δραστηριότητες μεγάλων αποστάσεων. Οι εξαρτήσεις αυτές είναι ροές που προέρχονται από βραχυκύκλωμα (loops). Ένα βραχυκύκλωμα μπορεί να επαναφέρει ένα trace σε μια αρχική δραστηριότητα.

3.1.3 Evolutionary Tree Miner

Το Evolutionary Tree Miner είναι ένας αλγόριθμος εξόρυξης γενετικών διαδικασιών ο οποίος παρουσιάστηκε ως ένας νέος αλγόριθμος που καθοδηγεί την ανακάλυψη διεργασιών με βάση τις τέσσερις μετρικές ποιότητας: a) της ικανότητας επανάληψης (Fitness), b) της ακρίβειας (Precision), c) της γενίκευσης (Generalization) και d) της απλότητας (Simplicity). Ο αλγόριθμος ETM στην παρούσα εργασία έχει υλοποιηθεί ως plug-in του ProM πλαίσιο εξόρυξης διεργασιών.

Το Input του αλγορίθμου ETM είναι ένα αρχείο καταγραφής (event log) συμβάντων που περιέχει την παρατηρούμενη συμπεριφορά μιας διεργασίας. Το πρώτο βήμα του αλγορίθμου είναι η τυχαία δημιουργία ενός πληθυσμού δέντρων διεργασιών. Για κάθε δέντρο, υπολογίζεται η συνολική βαθμολογία ποιότητας που βασίζεται σε ένα διαμορφώσιμο βάρος των τεσσάρων μετρικών ποιότητας η αλλιώς το κριτήριο για να σταματήσει η παραγωγή καινούργιων γενεών. Εάν έχει επιτευχθεί ένα από τα κριτήρια, όπως η ανακάλυψη ενός δένδρου διεργασίας με ένα συγκεκριμένο επιθυμητό fitness ή έχει παρέλθει ο μέγιστος αριθμός γενεών, τότε επιστρέφει το δέντρο διεργασίας με την

υψηλότερη βαθμολογία ποιότητας. Διαφορετικά, τα μοντέλα με τον τρέχοντα πληθυσμό των process trees θα πρέπει να βελτιωθούν. Τα καλύτερα μοντέλα του πληθυσμού (ελίτ), διατηρούνται και παραμένουν διαθέσιμα για να βελτιωθούν στις μελλοντικές γενιές. Ο υπόλοιπος πληθυσμός παραμένει και αυτός διαθέσιμος αλλά παρουσιάζει μικρότερες πιθανότητες τυχαίας επιλογής σε σχέση με τα καλύτερα μοντέλα. Αυτό δίνει μεγαλύτερη πιθανότητα επιλογής για αλλαγή στα καλύτερα μοντέλα. Υπάρχει πιθανότητα ένα μοντέλο να επιλεγεί πολλές φορές για να αλλάξει. Οι μετρικές ποιότητας υπολογίζονται ξανά για τη νέα γενιά μοντέλων και τα κριτήρια αξιολογούνται και πάλι. Ο κύκλος αυτός επαναλαμβάνεται μέχρις ότου ικανοποιηθεί ένα από τα κριτήρια που έχουμε προσδιορίσει.¹⁷

Τα τυχαία επιλεγμένα δέντρα επεξεργασίας μεταβάλλονται με τις εφαρμοσμένες λειτουργίες διασταύρωσης (crossover) και μετάλλαξης (mutation operations). Η λειτουργία crossover συνδυάζει δύο δέντρα διεργασιών για να δημιουργήσει δύο νέα δέντρα με χαρακτηριστικά από τα γονικά μοντέλα. Στη συνέχεια εφαρμόζεται η λειτουργία μετάλλαξης στα τυχαία επιλεγμένα δέντρα διεργασιών, συμπεριλαμβανομένων και αυτών που έχουν αλλάξει μέσω του crossover. Η λειτουργία της μετάλλαξης διαμορφώνει ένα δέντρο διεργασιών χρησιμοποιώντας τις παρακάτω μεταβολές:

1. Αντικατάσταση ενός δέντρου διεργασιών
2. Προσθήκη μιας δραστηριότητας
3. Αφαίρεση ενός υπό-δέντρου (sub-tree)
4. Αλλαγή μιας δραστηριότητας
5. Ολική αφαίρεση ενός δέντρου

3.1.4 Inductive Visual Miner

Το Inductive Miner περιέχει βελτιώσεις, σε σχέση με τους Alpha και Heuristic miner, που διευκολύνει την εξερεύνηση και την εξαγωγή πληροφοριών από ένα αρχείο καταγραφής συμβάντων. Παρέχει την ιδιαιτερότητα να αντιμετωπίζει τις σπάνιες συμπεριφορές όπως τον θόρυβο και μπορεί να χειριστεί με ευκολία τα μεγάλα αρχεία καταγραφής συμβάντων,

¹⁷ Maikel van Eck, BSc, (2013), "Alignment-based Process Model Repair and its Application to the Evolutionary Tree Miner"

εξασφαλίζοντας παράλληλα τις καλές μετρικές ποιότητας για τα μοντέλα διεργασίας που παράγει. Πρέπει να σημειωθεί ο Inductive Miner είναι σε θέση να ανακαλύψει διαδικασίες μάθησης παρόμοιες με εκείνες που χρησιμοποιούν οι εναλλακτικοί αλγόριθμοι που αναλύθηκαν. Ωστόσο η περίπτωση του Inductive Miner προσφέρεται για ανάλυση των δεδομένων καθώς η οπτικοποίηση που προσφέρει βοηθάει στην ερμηνεία των μοντέλων διεργασιών που παράγει. Οι αλγόριθμοι διερεύνησης διεργασιών συχνά οδηγούν σε πολύπλοκα μοντέλα, τα οποία είναι πολύ δύσκολο να διαβαστούν. Εντούτοις, ο Inductive Miner εστιάζει έντονα στην απλότητα και γενικά οδηγεί σε απλά μοντέλα.

Ο Inductive Visual Miner ανακαλύπτει αμέσως το αρχικό μοντέλο, υπολογίζει τις αποκλίσεις και τις παρουσιάζει στο χρήστη, χρησιμοποιώντας μια νέα απεικόνιση που επιτρέπει την κίνηση των event όπως τα έχει καταγράψει από το event log.

1. Κάνει έναν διαχωρισμό: μια κατανομή των δραστηριοτήτων στο αρχείο καταγραφής συμβάντων και ένα σχέδιο των δέντρων διεργασιών.
2. Χωρίζει το log σε μικρότερα sublogs.
3. Ανατρέχει σε κάθε sublog για να το χωρίσει σε μικρότερα και σταματά την αναδρομή όταν προκύψουν sublogs με μόνο ένα activity.
4. Χρησιμοποιεί την πληροφορία που έχει δομήσει για να φτιάξει την απεικόνιση.¹⁸

3.2 Μετρικές ποιότητας

Στο κεφάλαιο αυτό θα προσπαθήσουμε να προσδιορίσουμε τις μετρικές ποιότητας για την σύγκριση των αποτελεσμάτων των αλγορίθμων. Η προσπάθεια για την μέτρηση της ποιότητας είναι συνυφασμένη με τον έλεγχο συμμόρφωσης και την καταλληλότητα ενός προκαθορισμένου μοντέλου να αναπαράγει ένα Event Log. Ωστόσο, η καταλληλότητα είναι μόνο μία από τις τέσσερις μετρικές που χρησιμοποιούνται για την μέτρηση της ποιότητας.¹⁹

¹⁸ Alejandro Bogarín, Rebeca Cerezo, Cristóbal Romero, (2018) "Discovering learning processes using Inductive Miner: A case study with Learning Management Systems (LMSs)"

¹⁹ Carmona, J., van Dongen, B.F., Solti, A., Weidlich, M.(2018) "Conformance Checking Relating Processes and Models" Springer

Οι προκλήσεις που αντιμετωπίζουν οι αλγόριθμοι εξερεύνησης διεργασιών είναι η ατέλεια και ο θόρυβος. Η εξόρυξη διεργασιών είναι ένα σύνθετο έργο και πολλοί αλγόριθμοι έχουν προταθεί, για τον λόγο αυτό δεν είναι εύκολο να συγκρίνεις όλους αυτούς διαφορετικούς αλγορίθμους. Σε σύγκριση με τις κλασικές προκλήσεις της εξόρυξης δεδομένων, φαίνεται ότι υπάρχουν πολύ περισσότερες διαστάσεις τόσο από την άποψη της αναπαράστασης όσο και των κριτηρίων ποιότητας. Πολλές τεχνικές διερεύνησης διεργασιών αξιολογούνται χρησιμοποιώντας event logs από πραγματικά (real-life) δεδομένα και από πολλαπλά κριτήρια.

Παρόλο, που η καταλληλότητα (fitness) μπορεί να περιγράψει καλύτερα τον έλεγχο συμμόρφωσης, καλύπτει μόνο μία διάσταση. Η απλότητα, η ακρίβεια και η γενίκευση θα πρέπει να ληφθεί υπόψιν όταν αξιολογείται ένα μοντέλο ανακάλυψης. Αφήνοντας εκτός μια διάσταση μπορεί να οδηγήσει σε αποκεντρωμένα μοντέλα. Προφανώς, ένας αλγόριθμος διερεύνησης διεργασιών πρέπει να έχει την δυνατότητα να ερμηνεύσει μία διαφανόμενη συμπεριφορά. Κάνοντας μια επισκόπηση των μοντέλων είναι σημαντικό οι μετρικές που χρησιμοποιούμε να προσδιορίζουν την πολυπλοκότητα όπως και την ευκολία κατανόησης ενός μοντέλου διεργασιών. Επιπλέον μετρικές οι οποίες χρησιμοποιούμε είναι σημαντικό να λαμβάνονται υπόψιν είναι η δομή και η ομοιογένεια του μοντέλου.²⁰

Οι κυριότερες μετρικές προσδιορισμού της ποιότητας των αποτελεσμάτων της Εξόρυξης Διαδικασιών είναι οι εξής:

Replay Fitness

Η καταλληλότητα είναι η μετρική η οποία παρουσιάζει αν το εκάστοτε μοντέλο έχει την ικανότητα να αναπαράγει συμπεριφορά αντίστοιχη με αυτήν που έχει καταγραφεί στο event log από το οποίο το μοντέλο δημιουργήθηκε. Αν για παράδειγμα το μοντέλο δεν είναι σε θέση να αναπαράγει κάποιο trace τότε η καταλληλότητα μειώνεται.

Η καταλληλότητα υπολογίζεται ως εξής:

$$Q_{rf} = 1 - \frac{\text{cost for aligning model and event log}}{\text{Minimal cost to align arbitrary event log on model and vice versa}}$$

²⁰ Wil Van Der Aalst (2016) Process Mining: Data Science in Action 2nd edition Springer

Ο αριθμητής είναι η παράμετρος που εκφράζει το ποσοστό της αρνητικής συνέπειας για τις περιπτώσεις αλλαγών στο μοντέλο (εισαγωγή ή παράλειψη κάποιας διεργασίας). Ο παρονομαστής εκφράζει το ελάχιστο κόστος για την περίπτωση που το μοντέλο είναι τελείως διαφορετικό με αυτό που περιγράφεται στο event log (η χειρότερη περίπτωση). Η τιμή της καταλληλότητας του μοντέλου κυμαίνεται από την τιμή 0 έως την τιμή 1.²¹

Simplicity – Απλότητα

Η απλότητα ποσοτικοποιεί την πολυπλοκότητα του μοντέλου. Ο υπολογισμός της απλότητας προκύπτει από την σύγκριση του μεγέθους του μοντέλου που αναπαράχθηκε με τον αριθμό των δραστηριοτήτων από το event log. Η απλούστερη μορφή ενός μοντέλου είναι αυτή στην οποία κάθε δραστηριότητα εκτελείτε μόνο μία φορά.²²

Η απλότητα υπολογίζεται ως εξής:

$$Q_s = 1 - \frac{\#duplicate\ activities + \#missing\ activities}{\#nodes\ in\ process\ tree + \#event\ classes\ in\ event\ log}$$

Precision – Ακρίβεια

Ένα μοντέλο θεωρείται απόλυτα ακριβές όταν δεν έχει καθόλου escaping edges δηλαδή δραστηριότητες τις οποίες αναπαράγει το μοντέλο αλλά δεν είναι καταγεγραμμένες στον event log.²³

Η ακρίβεια υπολογίζεται ως εξής:

$$Q_s = 1 - \frac{\sum visited\ markings\ \#visits * \frac{outgoing\ edges - \#used\ edges}{outgoing\ edges}}{\#total\ marking\ visits\ over\ all\ markings}$$

Generalization – Γενίκευση

Το ποσό της γενίκευσης ενός μοντέλου εξαρτάται από την συχνότητα εκτέλεσης όλων των δραστηριοτήτων του. Αν για παράδειγμα κάποιες δραστηριότητες του μοντέλου δεν εκτελούνται συχνά τότε η γενίκευση δεν είναι

²¹ Fabian Rojas Blum, (2015), “Metrics in process discovery”

²² Fabian Rojas Blum, (2015), “Metrics in process discovery”

²³ Fabian Rojas Blum, (2015), “Metrics in process discovery”

η βέλτιστη αν όμως όλα τα μέρη του μοντέλου παρουσιάζουν καλή συχνότητα τότε είναι πολύ πιθανόν η γενίκευση να είναι καλή.²⁴

Η γενίκευση υπολογίζεται ως εξής:

$$Q_s = 1 - \frac{\sum \text{nodes}(\sqrt{\# \text{executions}}) - 1}{\text{nodes in tree}}$$

²⁴ Fabian Rojas Blum, (2015), "Metrics in process discovery"

4 Διεξαγωγή πειραμάτων

4.1 Εισαγωγή στην διαδικασία διεξαγωγής πειραμάτων

Στόχος της συγκεκριμένης μελέτης είναι να παραχθεί ένα σύνθετο μοντέλο διεργασιών και να εκτελεστεί με τους μεγαλύτερους δυνατούς αριθμούς επαναλήψεων. Στην συνέχεια τα παραγόμενα log αρχεία τύπου xes θα χρησιμοποιηθούν για να δοκιμαστούν οι τέσσερις αλγόριθμοι. Κάθε αλγόριθμος αντιστοιχεί σε ένα plugin του prom. Οι αριθμοί των επαναλήψεων για τους οποίους θα πραγματοποιηθούν τα πειράματα είναι 500, 1000 και 1500. Τα πειράματα εκτελέστηκαν σε προσωπικό υπολογιστή.

- **Επεξεργαστής – CPU:** AMD RYZEN 7 2700X 8-Core 3.7GHz
- **Μνήμη RAM:** 16GB

Plugins που θα χρησιμοποιηθούν για την δοκιμή των τεσσάρων αλγορίθμων

Για την διεξαγωγή των πειραμάτων θα χρησιμοποιηθούν plugins του Prom. Στο σκέλος του κάθε πειράματος που αφορά το Prom το αρχικό plugin του συγκεκριμένου σκέλους αντιστοιχεί με τον κάθε αλγόριθμο. Κάθε παραγόμενο αρχείο εάν δεν είναι ήδη Petri Net θα μετατρέπεται σε Petri Net με κάποιο αντίστοιχο Plugin. Στην συνέχεια θα γίνεται χρήση του Petri Net για να παραχθούν μέσω άλλων Plugins οι μετρικές ποιότητας. Για τον αλγόριθμο Alpha Miner θα γίνει χρήση του plugin “Alpha Miner”. Για τον αλγόριθμο Inductive Miner θα γίνει χρήση του Plugin “Mine with Inductive Visual Miner”. Για τον αλγόριθμο Heuristic Miner θα γίνει η χρήση του Plugin “Mine for a Heuristics Net with Heuristics Miner”. Για τον αλγόριθμο Evolutionary Tree Miner θα γίνει χρήση του Plugin “Mine a Process Tree with ETMd”.

Plugins για την μετατροπή των παραγόμενων αρχείων σε Petri Nets

Για την διεξαγωγή των μετρικών ποιότητας είναι απαραίτητα τα Petri Nets που αντιστοιχούν σε κάθε πείραμα. Για τον λόγο ότι κάποια plugins δεν παράγουν απευθείας Petri Nets έγινε σχετική έρευνα και επιλέχθηκαν τα παρακάτω Plugins: Για την μετατροπή του Process Tree που παράγεται από τον Evolutionary Tree Miner θα γίνει χρήση του plugin “Convert Process Tree

to Petri Net”. Για την μετατροπή του Heuristic Net που παράγεται από τον Heuristic Miner θα γίνει χρήση του “Convert Heuristic Net to Petri Net”.

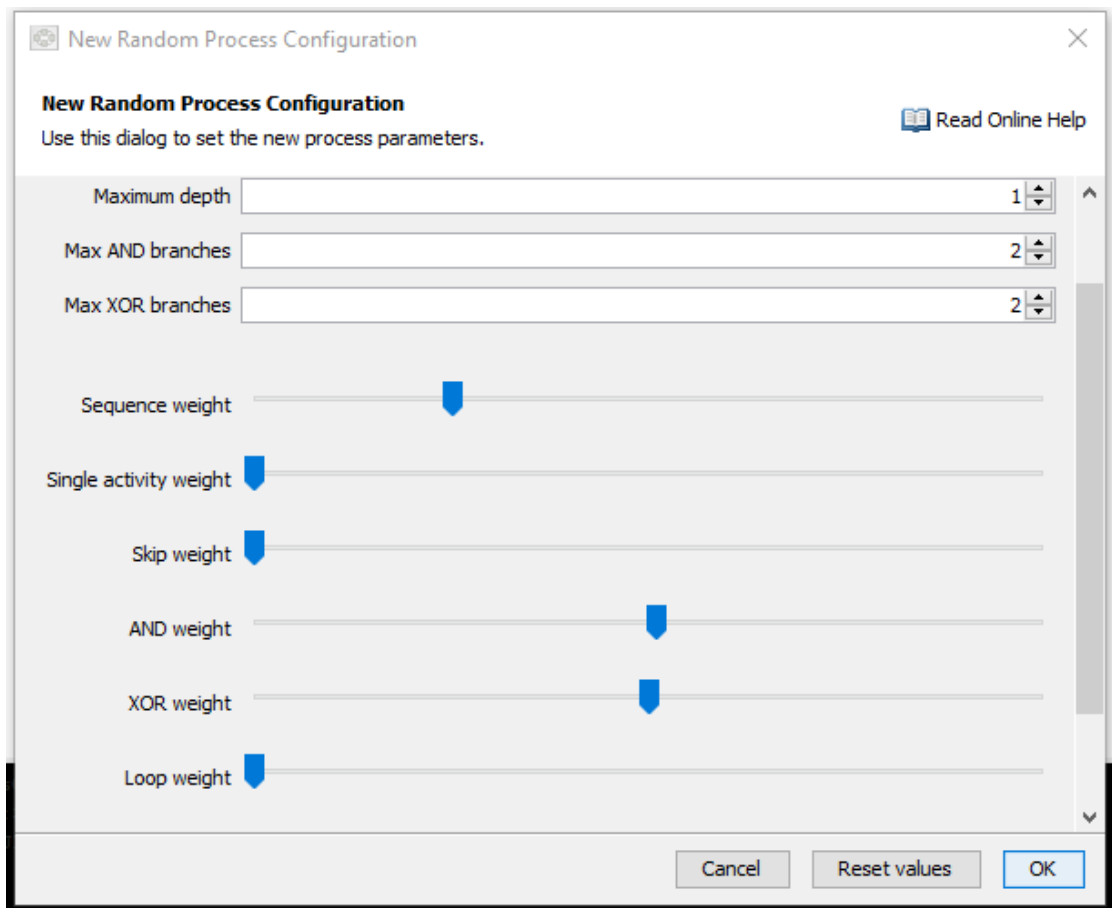
Plugins για διεξαγωγή μετρικών ποιότητας

Για την διεξαγωγή μετρικών ποιότητας επιλέχθηκαν τα Plugins “Compute Projected Fitness and Precision” για το Fitness και για το Precision και το Generalization επιλέχθηκε το “Measure Precision/ Generalization”. Το plugin “Compute Projected Fitness and Precision” για να επεξεργαστεί το εκάστοτε Petri Net θα πρέπει να γίνει πρώτα μία προεργασία και μετατροπή του με το plugin “Convert Petri Net to Accepting Petri Net”. Επίσης αντίστοιχα πρέπει να γίνει προεργασία για το “Measure Precision/ Generalization” που λαμβάνει σαν είσοδο ένα aligned replay που προκύπτει από το παραγόμενο petri net και το σχετικό xes log με την χρήση του Plugin “Align Log and Model for Repair”.

4.2 Παραγωγή σύνθετου μοντέλου με την χρήση του λογισμικού Plg2

Στο PLG2, η εσωτερική δομή ενός μοντέλου διεργασίας εισάγεται διαισθητικά από τον ορισμό ενός μοντέλου διεργασίας BPMN. Μια διαδικασία είναι ουσιαστικά μια συνάθροιση των συνιστωσών. Κάθε στοιχείο μπορεί να είναι είτε αντικείμενο ροής, ακολουθία ή αντικείμενο δεδομένων.

Η δημιουργία τυχαίων διεργασιών βασίζεται σε μερικούς τρόπους ελέγχου της ροής εργασίας. Στην εικόνα 6 παρουσιάζονται οι τρόποι ροής χρησιμοποιήσαμε για την παραγωγή ενός του σύνθετου μοντέλου κάνοντας χρήση των παραμέτρων Sequence Weight , Xor Weight και And Weight.



Εικόνα 6: Παραγωγή σύνθετου μοντέλου με την χρήση του PLG2

Τα αντικείμενα δεδομένων παράγονται τυχαία και τυχαία συνδέονται με τις δραστηριότητες. Το τυχαίο μοντέλο που επιλέχθηκε έχει βάθος τριών δραστηριοτήτων και το άθροισμα των δραστηριοτήτων ανέρχεται στα 8



Εικόνα 7: Σύνθετο μοντέλο BPMN για την διεξαγωγή των πειραμάτων

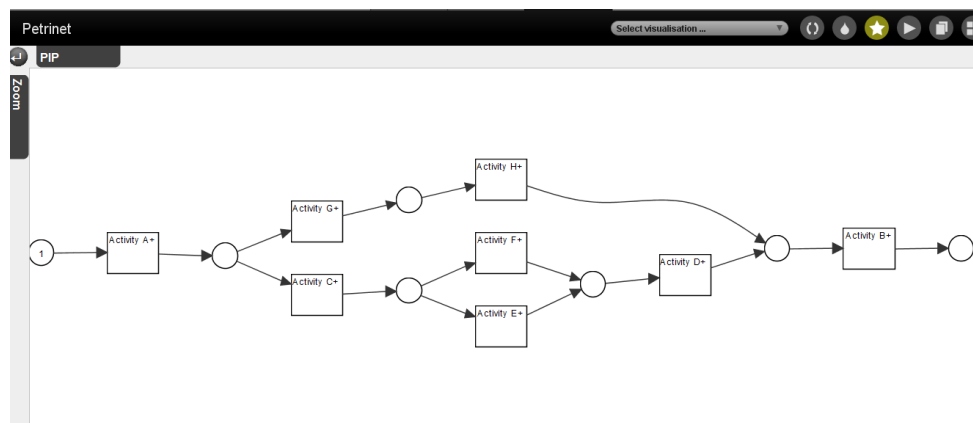
4.3 Πείραμα με 500 εκτελέσεις

Στο πρώτο σκέλος του πρώτου πειράματος παράχθηκε log αρχείο τύπου xes με 500 εκτελέσεις χωρίς θόρυβο. Η εκτέλεση του BPMN μοντέλου χωρίς θόρυβο θα επαναληφθεί και στα επόμενα πειράματα για να γίνουν πιο ευδιάκριτες οι επιπτώσεις του θορύβου που θα προστίθενται στην συνέχεια. Τα

πρώτα συμπεράσματα των πειραμάτων προκύπτουν από την γενική εικόνα των δικτύων που δίνει η εκτέλεση των σχετικών plugin του prom π.χ “Heuristic Net, Petri Net” κ.λ.π.. Αφού ολοκληρωθεί η δημιουργία των δικτύων από την εκτέλεση των αλγορίθμων θα συνεχιστεί η μελέτη με την διεξαγωγή των μετρικών ποιότητας και αφού ολοκληρωθούν οι παραπάνω διαδικασίες θα γίνει μία συνολική μελέτη σχετικά με όλα τα σκέλη των τριών πειραμάτων.

Εκτέλεση Alpha Miner για log με 500 εκτελέσεις χωρίς θόρυβο

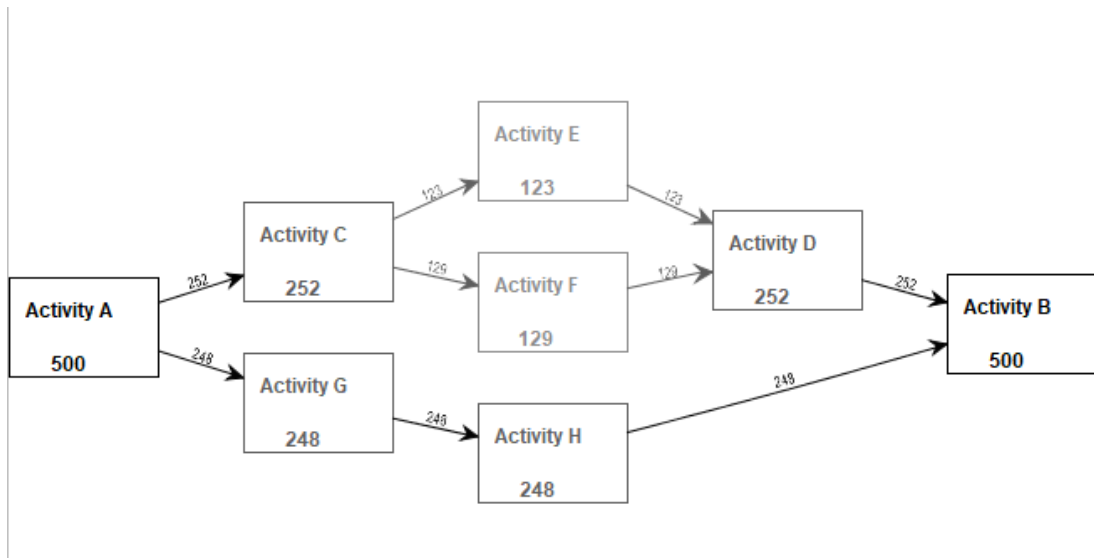
Με την εκτέλεση του Alpha Miner μέσω του Prom plugin “Alpha Miner” είναι αρκετά ευδιάκριτο ότι δεν υπάρχουν καθόλου σφάλματα και οι δραστηριότητες του Petri Net αντιστοιχούν απόλυτα με αυτές του BPMN.



Εικόνα 8: Petri Net από το Alpha Miner Plugin για το πείραμα με τις 500 εκτελέσεις και χωρίς θόρυβο

Εκτέλεση Heuristic Miner για log με 500 εκτελέσεις χωρίς θόρυβο

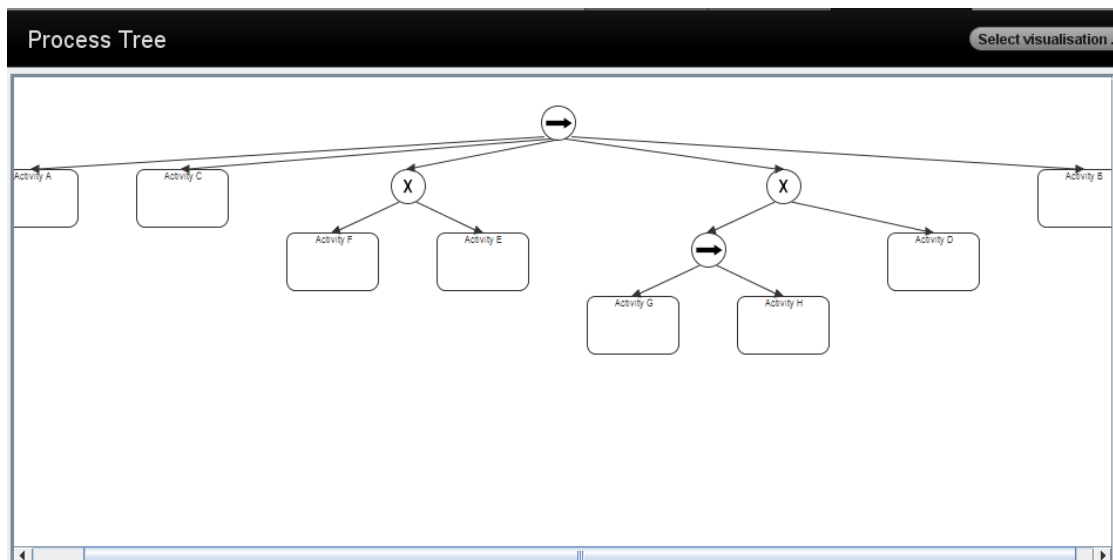
Με την εκτέλεση του Heuristic Miner για το πρώτο σκέλος με τις 500 εκτελέσεις και χωρίς καθόλου θόρυβο προέκυψε το αντίστοιχο Heuristic Net στο οποίο παρατηρείται ότι οι δραστηριότητες που βρίσκονται σε σειρά εκτελέστηκαν σε κάθε επανάληψη ενώ παρατηρείται ότι οι εκτελέσεις των παράλληλων δραστηριοτήτων είναι σχεδόν ισόποσα μοιρασμένες.



Εικόνα 9: Heuristic Net από το Mine a Heuristic Net with the Heuristic Miner Plugin για το πείραμα με τις 500 εκτελέσεις και χωρίς θόρυβο

Εκτέλεση Evolutionary tree Miner για log με 500 εκτελέσεις χωρίς θόρυβο

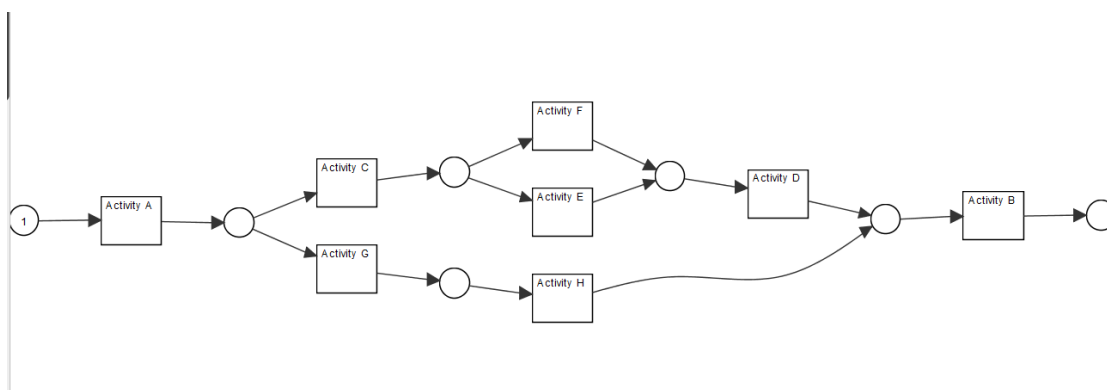
Στο τρέχον σκέλος του πρώτου πειράματος παράχθηκε ένα Process Tree με την χρήση του Evolutionary Tree Mine μέσω του plugin “Mine a Process Tree with ETMd”. Είναι αρκετά ευδιάκριτο ότι δεν υπάρχουν καθόλου σφάλματα και οι δραστηριότητες του Petri Net αντιστοιχούν απόλυτα με αυτές του BPMN



Εικόνα 10: Process Tree από το Mine a Process Tree with ETMd Plugin για το πείραμα με τις 500 εκτελέσεις και χωρίς θόρυβο

Εκτέλεση Inductive Miner για log με 500 εκτελέσεις χωρίς θόρυβο

Με την εκτέλεση του Inductive Miner προκύπτει ένα Petri Net το οποίο επίσης δεν παρουσιάζει καθόλου σφάλματα και οι δραστηριότητες συμφωνούν με το BPMN του πειράματος. Για την δημιουργία του συγκεκριμένου Petri Net έγινε χρήση του Mine Petri Net With Inductive Miner.



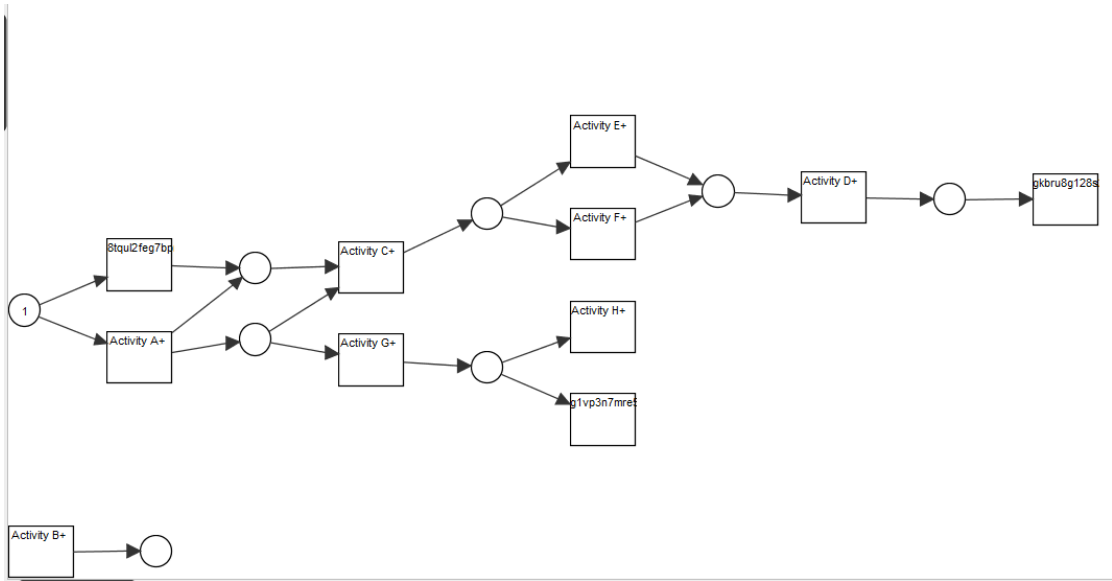
Εικόνα 11: Petri Net από το Mine Petri Net With Inductive Miner Plugin για το πείραμα με τις 500 εκτελέσεις και χωρίς θόρυβο

Δεύτερο σκέλος του πειράματος με 500 εκτελέσεις και προσθήκη θορύβου μόνο στα Activity Names

Στο συγκεκριμένο σκέλος του πειράματος έγινε προσθήκη θορύβου στα activity name του BPMN μοντέλου και εκτελέστηκε 500 φορές. Από αυτό το σκέλος και στην συνέχεια του πειράματος θα αρχίσουν να είναι εμφανείς οι επιπτώσεις του θορύβου σε ένα τέτοιο μοντέλο

Εκτέλεση Alpha Miner για log με 500 εκτελέσεις με θόρυβο στα activity names

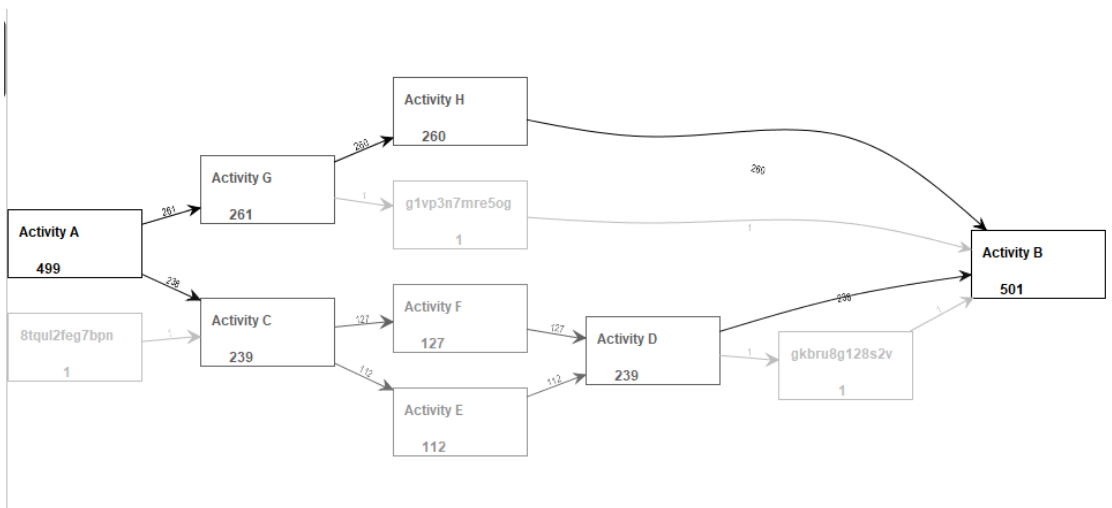
Για το συγκεκριμένο σκέλος του πειράματος και με την χρήση του Alpha Miner plugin του Prom προέκυψε το Petri Net της εικόνας 12. Είναι προφανής η παρουσία λίγων σφαλμάτων αλλά και η αλλαγή του σχήματος του Petri Net με την δραστηριότητα D να έχει αποκολληθεί από το υπόλοιπο Petri Net. Επίσης μπορεί κανείς να διαπιστώσει τις επιπτώσεις του θορύβου να συγκρίνει το τρέχον Petri Net με το τέλειο Petri Net από το προηγούμενο σκέλος του πειράματος.



Εικόνα 12: Petri Net από το Alpha Miner Plugin για το πείραμα με τις 500 εκτελέσεις και θόρυβο στα activity names

Εκτέλεση Heuristic Miner για log με 500 εκτελέσεις με θόρυβο στα activity names

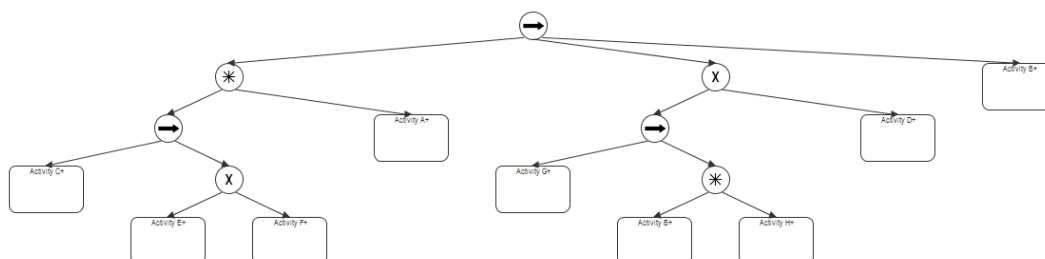
Με την εκτέλεση του Heuristic Miner προκύπτει ένα Heuristic Net στο οποίο προφανής η παρουσία λίγων σφαλμάτων αλλά και η αλλαγή του σχήματος του. Επίσης παρατηρείται μείωση των συχνοτήτων στις παράλληλες δραστηριότητες του δικτύου. Εικόνα 13



Εικόνα 13: Heuristic Net από το Mine a Heuristic Net with the Heuristic Miner Plugin για το πείραμα με τις 500 εκτελέσεις με θόρυβο στα activity names

Εκτέλεση Evolutionary tree Miner για log με 500 εκτελέσεις με θόρυβο στα activity names

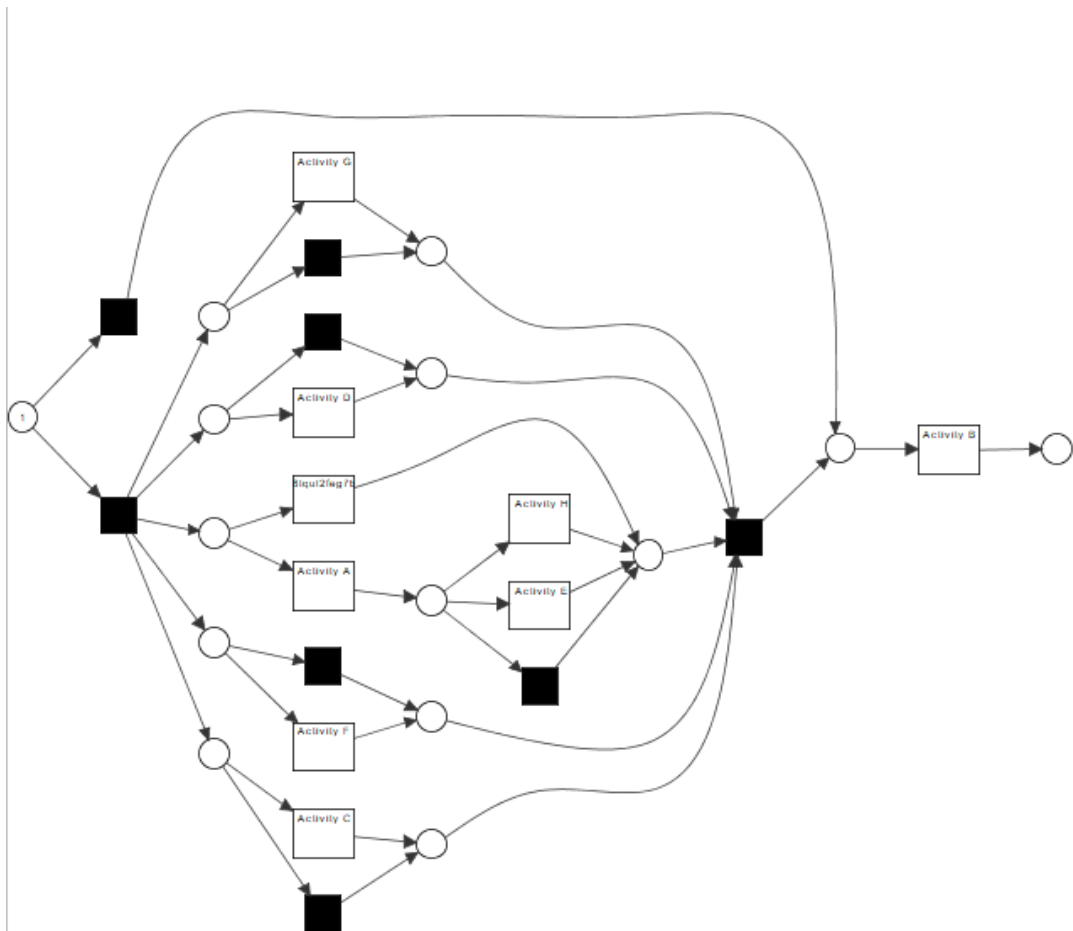
Στο τρέχον σκέλος του πειράματος παράχθηκε ένα Process Tree με την χρήση του Evolutionary Tree Mine μέσω του plugin “Mine a Process Tree with ETMd”. Είναι αρκετά προφανές ότι έχει αλλάξει η σειρά των δραστηριοτήτων και ότι η δραστηριότητα B έχει διαχωριστεί σε δύο σκέλη Εικόνας 14



Εικόνα 14: Process Tree από το Mine a Process Tree with ETMd Plugin για το πείραμα με τις 500 εκτελέσεις με θόρυβο στα activity names

Εκτέλεση Inductive Miner για log με 500 εκτελέσεις με θόρυβο στα activity names

Με την εκτέλεση του Inductive Miner προκύπτει ένα Petri Net στο οποίο παρουσιάζονται σφάλματα που καταστούν περισσότερο πολύπλοκη την δομή του δικτύου και κατά συνέπεια τις σχέσεις των δραστηριοτήτων. Εικόνα 15



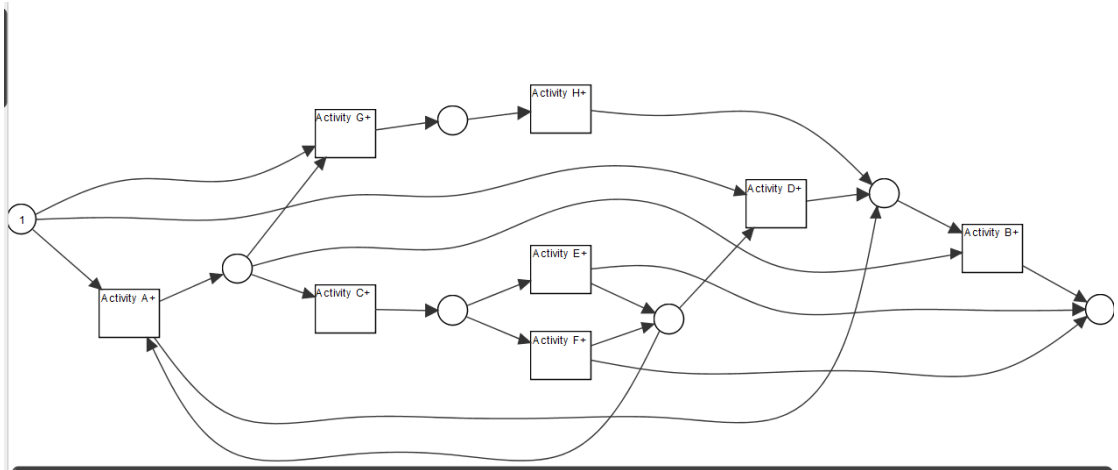
Εικόνα 15: Petri Net από το Mine Petri Net With Inductive Miner Plugin για το πείραμα με τις 500 εκτελέσεις και θόρυβο στα activity names

Τρίτο σκέλος του πειράματος με 500 εκτελέσεις και προσθήκη θορύβου μόνο στα control flows

Στο συγκεκριμένο σκέλος του πειράματος έγινε προσθήκη θορύβου στα control flows του BPMN μοντέλου.

Εκτέλεση Alpha Miner για log με 500 εκτελέσεις με θόρυβο στα control flows

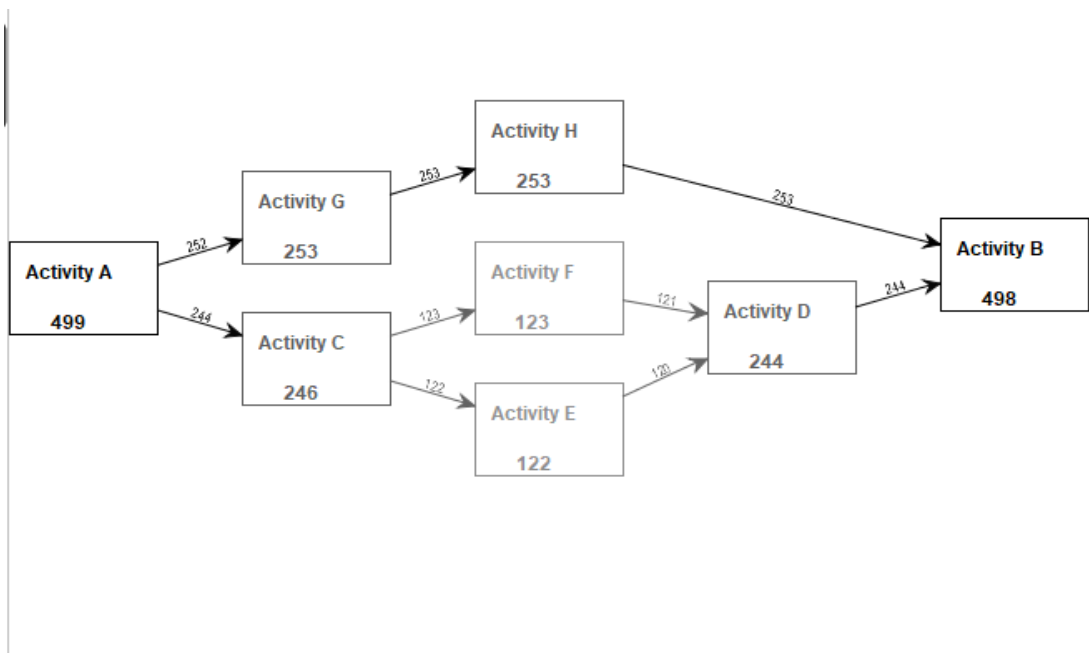
Με την εκτέλεση του Alpha Miner, μέσω του Prom plugin “Alpha Miner”, είναι αρκετά προφανές ότι έχουν αλλάξει οι σχέσεις των δραστηριοτήτων του Petri Net. Στο συγκεκριμένο Petri Net βλέπουμε ότι πέρα από τις αλλαγές στις σχέσεις των δραστηριοτήτων δεν έχουν παρουσιαστεί καθόλου σφάλματα Εικόνες 16.



Εικόνα 16: Petri Net από το Alpha Miner Plugin για το πείραμα με τις 500 με θόρυβο στα control flows

Εκτέλεση Heuristic Miner για log με 500 εκτελέσεις με θόρυβο στα control flows

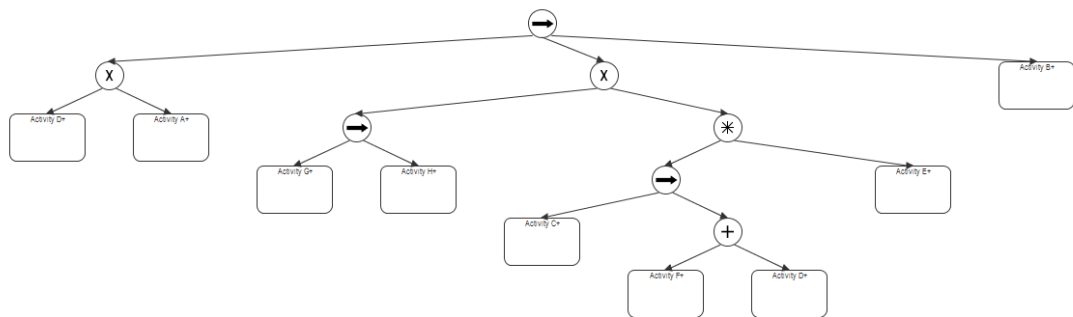
Από το Heuristic Net που προέκυψε σε αυτό το σκέλος του πειράματος φαίνεται πως ο θόρυβος στα control flows του δικτύου δεν επηρέασε την ποιότητα του συγκεκριμένου δικτύου. Φαίνεται πως οι δραστηριότητες σε σειρά έχουν πολύ καλή συχνότητα και πως στις παράλληλες δραστηριότητες οι συχνότητες έχουν διαχωριστεί σχεδόν ισόποσα. Το τρέχον Heuristic Net είναι αρκετά παρόμοιο με το heuristic net από το σκέλος του πειράματος με τις 500 επαναλήψεις χωρίς θόρυβο.



Εικόνα 17: Heuristic Net από το Mine a Heuristic Net with the Heuristic Miner Plugin για το πείραμα με τις 500 εκτελέσεις με θόρυβο στα control flows

Εκτέλεση Evolutionary tree Miner για log με 500 εκτελέσεις με θόρυβο στα control flows

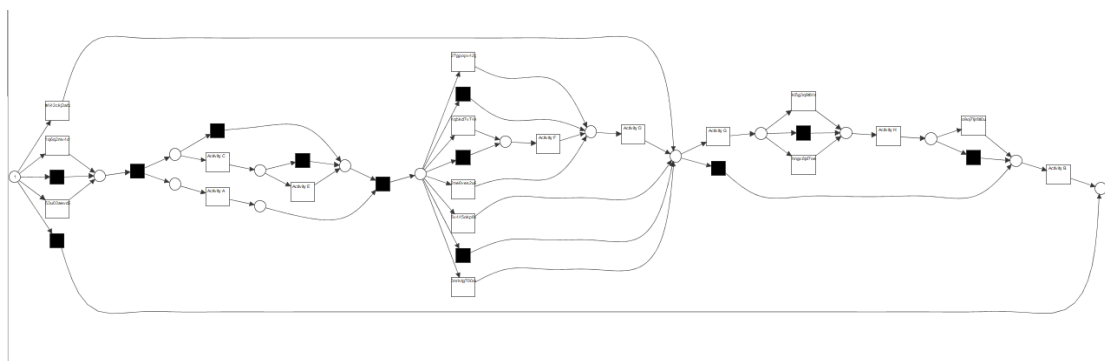
Στο τρέχον σκέλος του πειράματος παράχθηκε ένα Process Tree με την χρήση του Evolutionary Tree Mine μέσω του plugin “Mine a Process Tree with ETMd”. Το τρέχον Process Tree είναι αρκετά παρόμοιο με το Process Tree από το σκέλος του πειράματος με τις 500 επαναλήψεις χωρίς θόρυβο.



Εικόνα 18: Process Tree από το Mine a Process Tree with ETMd Plugin για το πείραμα με τις 500 εκτελέσεις με θόρυβο στα control flows

Εκτέλεση Inductive Miner για log με 500 εκτελέσεις και θόρυβο στα control flows

Με την εκτέλεση του Inductive Miner προκύπτει ένα Petri Net στο οποίο παρουσιάζονται σφάλματα που καταστούν περισσότερο πολύπλοκη την δομή του δικτύου και κατά συνέπεια τις σχέσεις των δραστηριοτήτων. Παρατηρείται ότι το τρέχον petri net είναι ακόμη πιο πολύπλοκο από αυτό που προέκυψε στο προηγούμενο πείραμα με την χρήση του Inductive Miner.



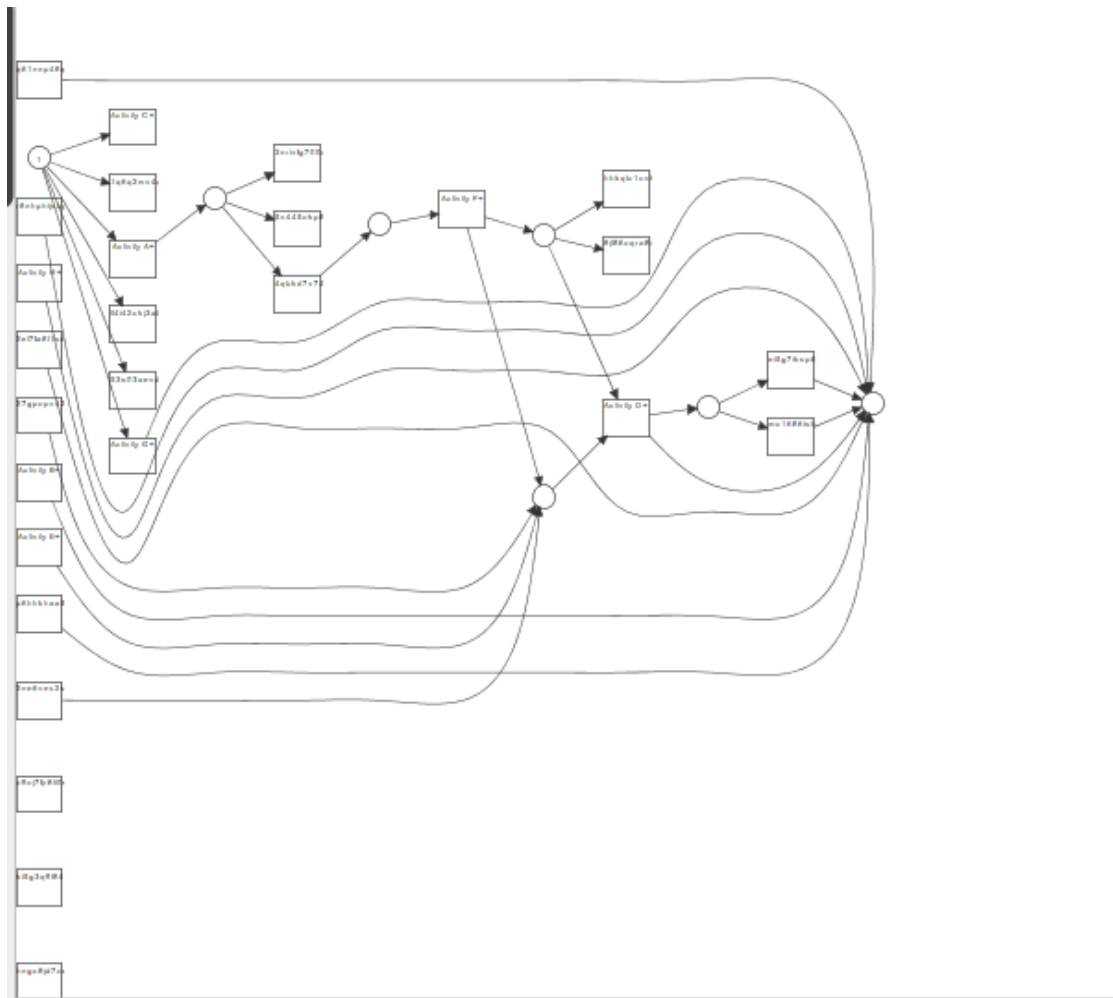
Εικόνα 19: Petri Net από το Mine Petri Net With Inductive Miner Plugin για το πείραμα με τις 500 εκτελέσεις και θόρυβο στα control flows

Τέταρτο σκέλος του πειράματος με 500 εκτελέσεις και προσθήκη complete noise

Στο συγκεκριμένο σκέλος του πειράματος έγινε προσθήκη θορύβου complete noise δηλαδή σε όλο το φάσμα του BPMN μοντέλου και έγινε εκτέλεση του μοντέλου 500 φορές.

Εκτέλεση Alpha Miner για log με 500 εκτελέσεις με θόρυβο σε όλο το φάσμα του μοντέλου

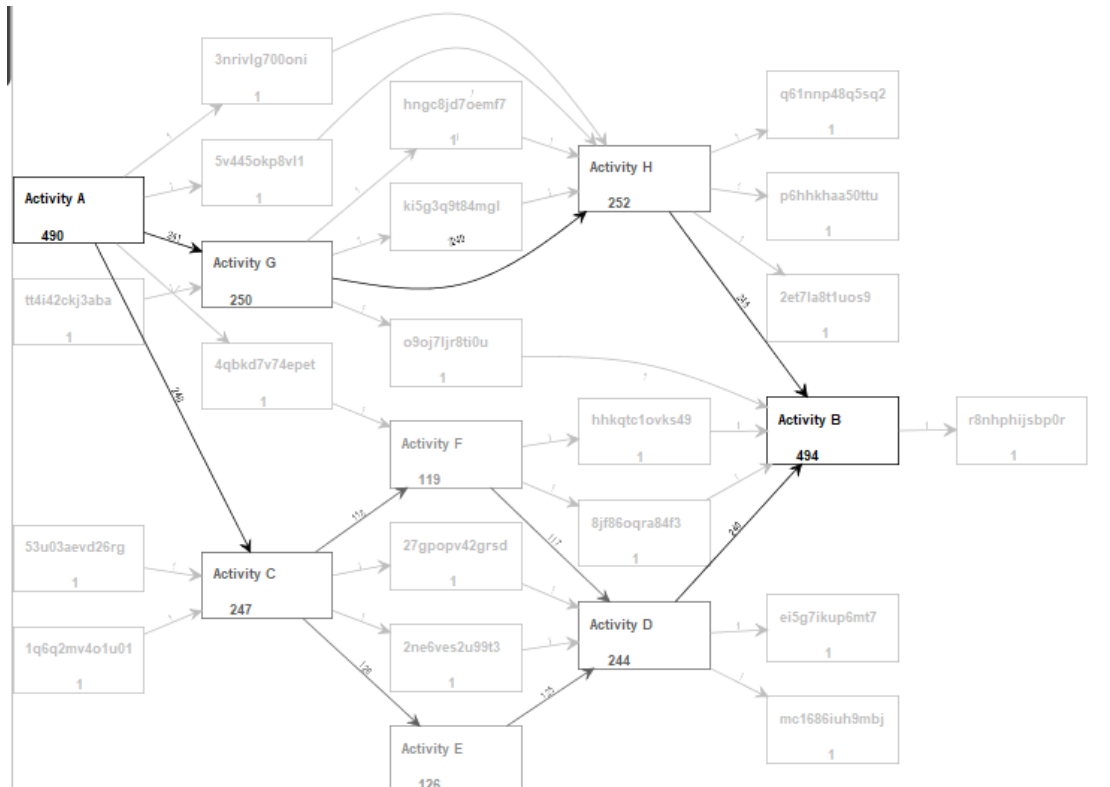
Στο συγκεκριμένο σκέλος του πειράματος έγινε εκτέλεση του Alpha Miner με θόρυβο σε όλο το φάσμα του μοντέλου και προέκυψε ένα petri net με πολλά σφάλματα που το καταστούν αρκετά πολύπλοκο.



Εικόνα 20: Petri Net από το Alpha Miner Plugin για το πείραμα με τις 500 εκτελέσεις με πλήρη θόρυβο

Εκτέλεση Heuristic Miner για log με 500 εκτελέσεις με θόρυβο στα control flows

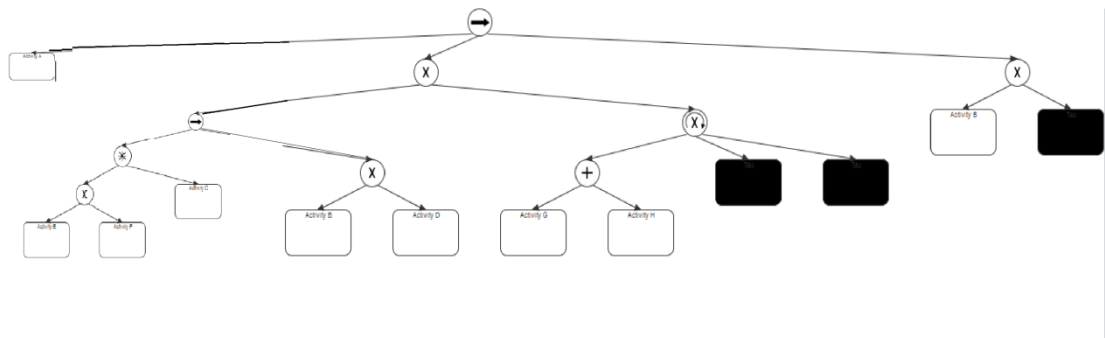
Από το Heuristic Net που προέκυψε σε αυτό το σκέλος του πειράματος φαίνεται πως ο θόρυβος σε όλο το φάσμα του πειράματος μείωσε πάρα πολύ την ποιότητα του heuristic net με αποτέλεσμα αν είναι αρκετά πολύπλοκο και δυσνόητο. Είναι εμφανής η παρουσία πολλών σφαλμάτων αλλά και η μείωση των συχνοτήτων σε όλες τις δραστηριότητες σε σχέση με τα προηγούμενα Heuristic Nets.



Εικόνα 21: Heuristic Net από το Mine a Heuristic Net with the Heuristic Miner Plugin για το πείραμα με τις 500 εκτελέσεις με θόρυβο σε όλο το φάσμα του μοντέλου

Εκτέλεση Evolutionary tree Miner για log με 500 εκτελέσεις με πλήρη θόρυβο

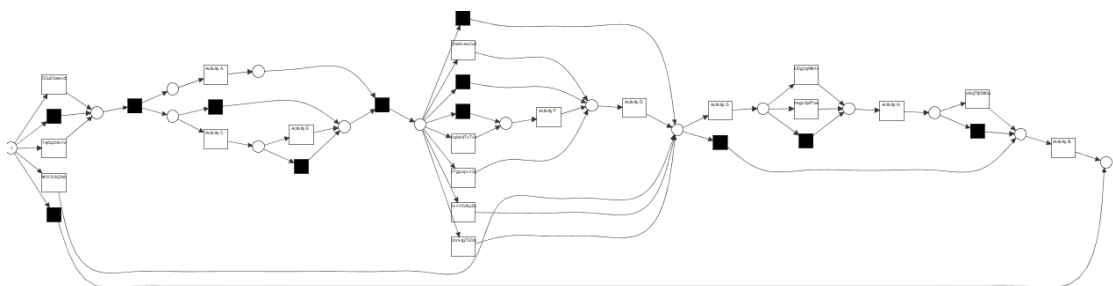
Με το process tree που προκύπτει από την εκτέλεση του Evolutionary tree miner είναι αρκετά προφανές ότι ο θόρυβος σε όλο το φάσμα του μοντέλου προκάλεσε σφάλματα αλλά επίσης επηρέασε την δομή του process tree.



Εικόνα 22: Process Tree από το Mine a Process Tree with ETMd Plugin για το πείραμα με τις 500 εκτελέσεις με πλήρη θόρυβο

Εκτέλεση Inductive Miner για log με 500 εκτελέσεις πλήρη θόρυβο

Με την εκτέλεση του Inductive Miner προκύπτει ένα Petri Net στο οποίο παρουσιάζονται σφάλματα που καταστούν ακόμη περισσότερο πολύπλοκη την δομή του δικτύου και κατά συνέπεια τις σχέσεις των δραστηριοτήτων. Παρατηρείται ότι το τρέχον petri net είναι ακόμη πιο πολύπλοκο από αυτό που προέκυψε στο προηγούμενο πείραμα με την χρήση του Inductive Miner.



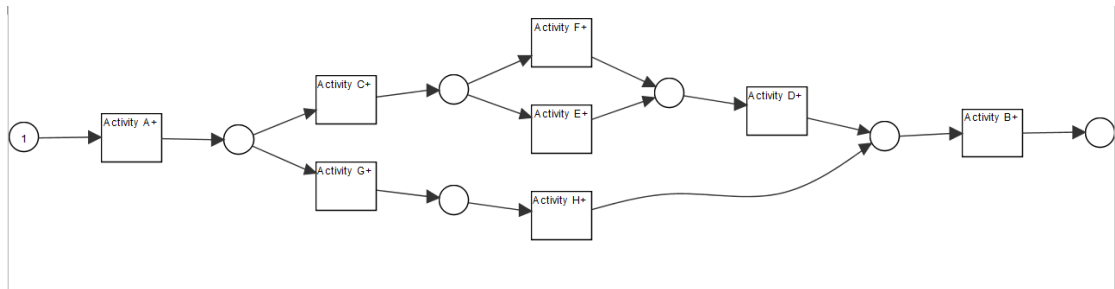
Εικόνα 23: Petri Net από το Mine Petri Net With Inductive Miner Plugin για το πείραμα με τις 500 εκτελέσεις και θόρυβο με πλήρη θόρυβο

4.4 Πείραμα με 1000 εκτελέσεις

Στο τρέχον πείραμα έγινε εκτέλεση του BPMN μοντέλου για 1000 φορές με την χρήση του λογισμικού PLG2. Με την χρήση του log αρχείου τύπου xes που προέκυψε θα επαναληφθεί η εκτέλεση των τεσσάρων αλγορίθμων με τα αντίστοιχα Prom Plugins.

Εκτέλεση Alpha Miner για log με 1000 εκτελέσεις χωρίς προσθήκη θορύβου

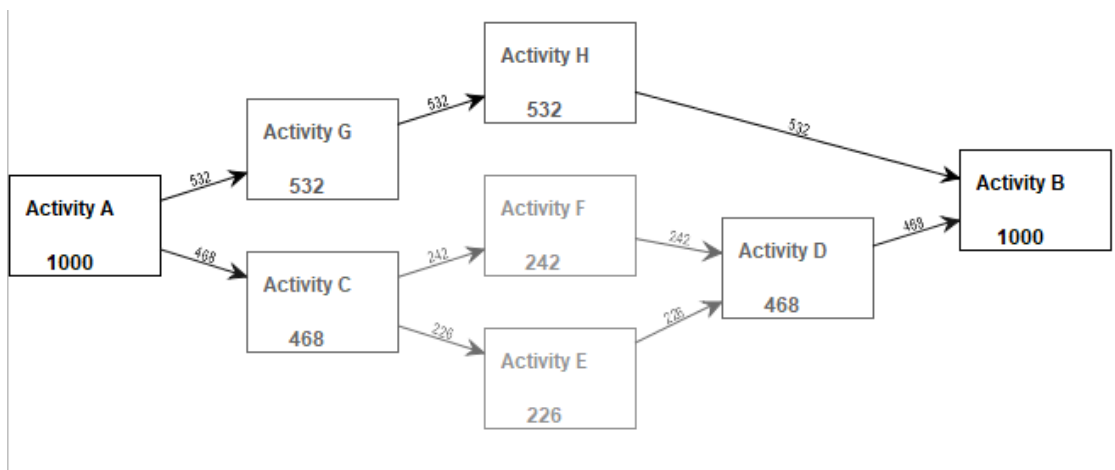
Σε αυτό το σκέλος του πειράματος με την εκτέλεση του Alpha Miner προκύπτει ένα τέλειο ποιοτικά Petri Net στο οποίο δεν παρατηρούνται καθόλου σφάλματα και η εικόνα του θυμίζει το BPMN μοντέλο αλλά και το αντίστοιχο Petri Net του αντίστοιχου προηγούμενου πειράματος με τις 500 εκτελέσεις.



Εικόνα 24: Petri Net από το Alpha Miner Plugin για το πείραμα με τις 1000 εκτελέσεις και χωρίς θόρυβο

Εκτέλεση Heuristic Miner για log με 1000 εκτελέσεις χωρίς θόρυβο

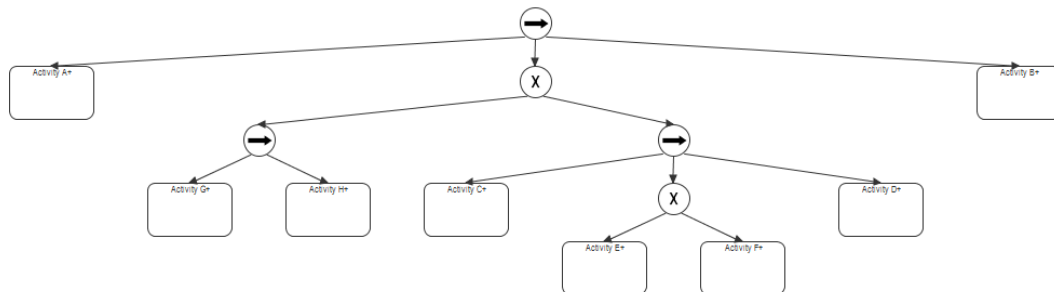
Από την εκτέλεση του Heuristic Miner με την Χρήση του Plugin Mine for a Heuristics Net using Heuristics Miner. Φαίνεται πως οι δραστηριότητες σε σειρά έχουν πολύ καλή συχνότητα και πως στις παράλληλες δραστηριότητες οι συχνότητες έχουν διαχωριστεί σχεδόν ισόποσα. Το τρέχον Heuristic Net είναι αρκετά παρόμοιο με το heuristic net από το σκέλος του πειράματος με τις 500 επαναλήψεις χωρίς θόρυβο.



Εικόνα 25: Heuristic Net από το Mine a Heuristic Net with the Heuristic Miner Plugin για το πείραμα με τις 1000 εκτελέσεις χωρίς θόρυβο

Εκτέλεση Evolutionary tree Miner για log 1000 εκτελέσεις χωρίς θόρυβο

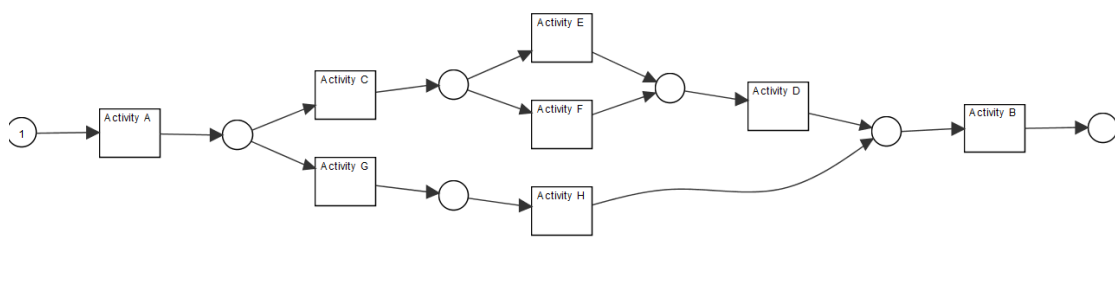
Στο τρέχον σκέλος του πειράματος παράχθηκε ένα Process Tree με την χρήση του Evolutionary Tree Mine μέσω του plugin “Mine a Process Tree with ETMd”. Το τρέχον Process Tree είναι άριστο ποιοτικά και ίδιο με το Process Tree από το σκέλος του πειράματος με τις 500 επαναλήψεις χωρίς θόρυβο.



Εικόνα 26: Process Tree από το Mine a Process Tree with ETMd Plugin για το πείραμα με τις 1000 επαναλήψεις

Εκτέλεση Inductive Miner για log με 1000 εκτελέσεις χωρίς θόρυβο

Σε αυτό το σκέλος του πειράματος με την εκτέλεση του Inductive Miner προκύπτει ένα τέλειο ποιοτικά Petri Net στο οποίο δεν παρατηρούνται καθόλου σφάλματα και η εικόνα του θυμίζει το BPMN μοντέλο αλλά και το αντίστοιχο Petri Net του προηγούμενου πειράματος.

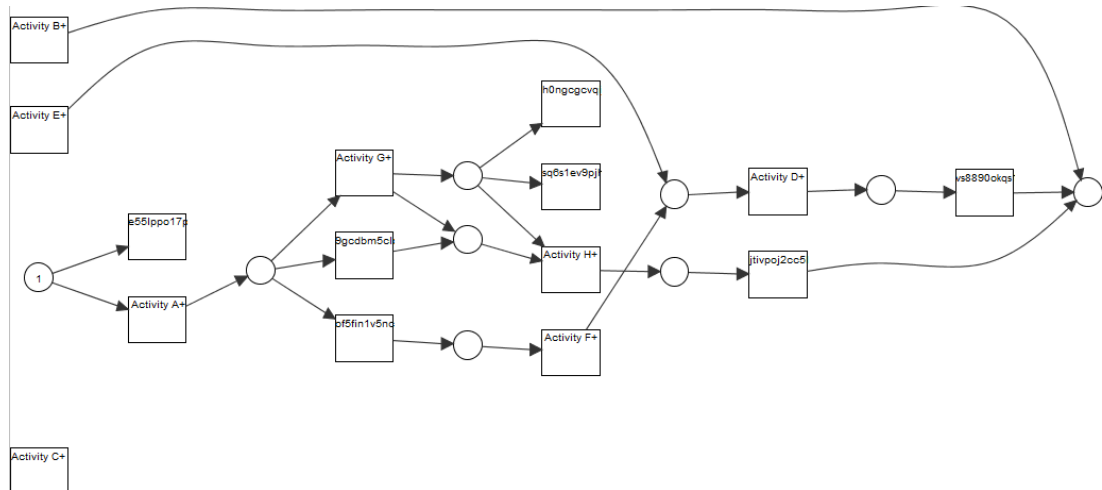


Εικόνα 27: Petri Net από το Mine Petri Net With Inductive Miner Plugin για το πείραμα με τις 1000 εκτελέσεις χωρίς θόρυβο

Εκτέλεση Alpha Miner για log με 1000 εκτελέσεις με προσθήκη θορύβου στα activity names

Για το συγκεκριμένο σκέλος του πειράματος και με την χρήση του Alpha Miner plugin του Prom προέκυψε το Petri Net της εικόνας 28. Είναι προφανής η παρουσία λίγων σφαλμάτων αλλά και η αλλαγή του σχήματος του Petri Net με τις δραστηριότητες C, B και E να έχουν αποκολληθεί από το υπόλοιπο Petri

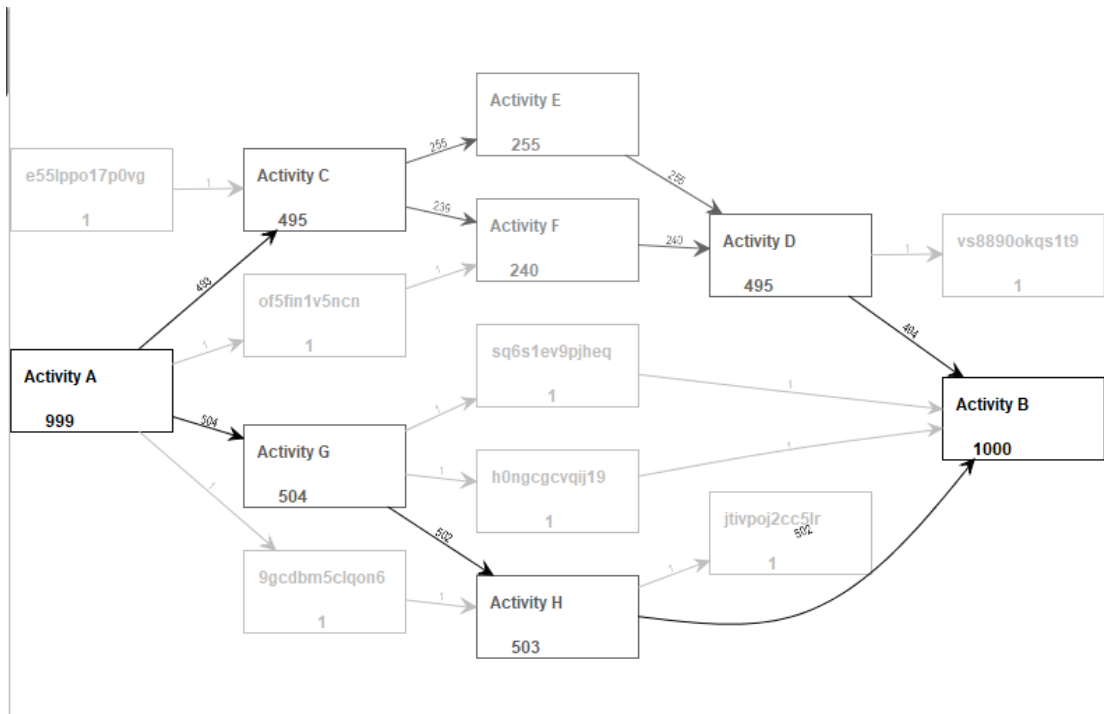
Net. Επίσης μπορεί κανείς να διαπιστώσει τις επιπτώσεις του θορύβου να συγκρίνει το τρέχον Petri Net με το τέλειο Petri Net από το προηγούμενο σκέλος του πειράματος στην εικόνα 24.



Εικόνα 28: Petri Net από το Alpha Miner Plugin για το πείραμα με τις 1000 εκτελέσεις με θόρυβο στα activity names

Εκτέλεση Heuristic Miner για log με 1000 με προσθήκη θορύβου στα activity names

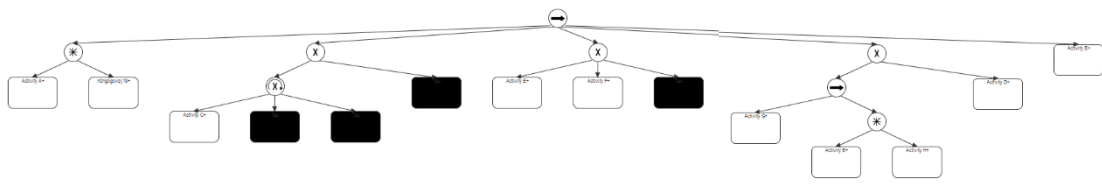
Από το Heuristic Net που προέκυψε σε αυτό το σκέλος του πειράματος φαίνεται πως ο θόρυβος στα activity name σε συνδυασμό με τις 1000 εκτελέσεις μείωσε πάρα πολύ την ποιότητα του heuristic net με αποτέλεσμα να είναι αρκετά πολύπλοκο και δυσνόητο. Είναι εμφανής η παρουσία πολλών σφαλμάτων αλλά και η μείωση των συχνοτήτων σε όλες τις δραστηριότητες.



Εικόνα 29: Heuristic Net από το Mine Heuristics Net with Heuristic Miner Plugin για το πείραμα με τις 1000 εκτελέσεις με θόρυβο στα activity names

Εκτέλεση Evolutionary tree Miner για log με 1000 εκτελέσεις με θόρυβο στα activity names

Με το process tree που προκύπτει από την εκτέλεση του Evolutionary tree miner είναι αρκετά προφανές ότι ο θόρυβος στα activity names του μοντέλου σε συνδυασμό με τις 1000 εκτελέσεις προκάλεσαν σφάλματα αλλά επίσης επηρέασε την δομή του process tree.

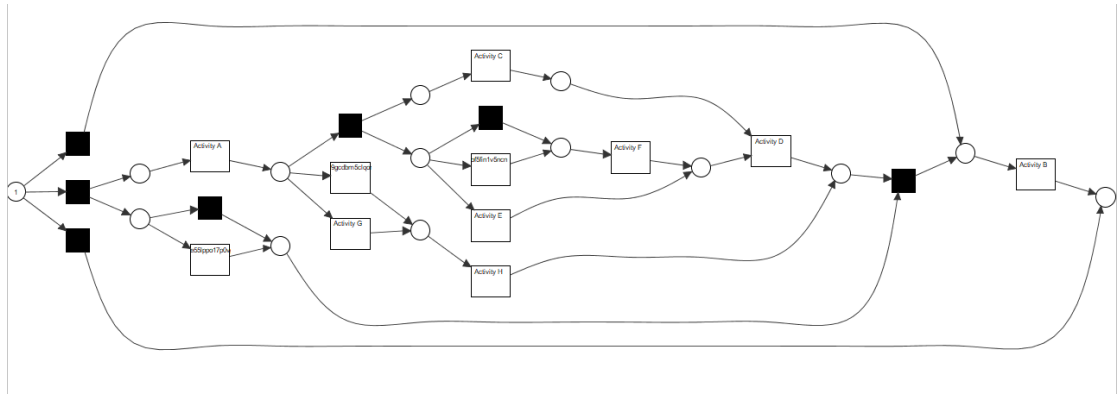


Εικόνα 30: Process Tree από το Mine a Process Tree with ETMd Plugin για το πείραμα με τις 1000 εκτελέσεις με θόρυβο στα activity names

Εκτέλεση Inductive Miner για log με 1000 εκτελέσεις με θόρυβο στα activity names

Με την εκτέλεση του Inductive Miner προκύπτει ένα Petri Net στο οποίο παρουσιάζονται σφάλματα που καταστρέφουν πολύπλοκη την δομή του δικτύου και κατά συνέπεια τις σχέσεις των δραστηριοτήτων. Παρατηρείται ότι το τρέχον

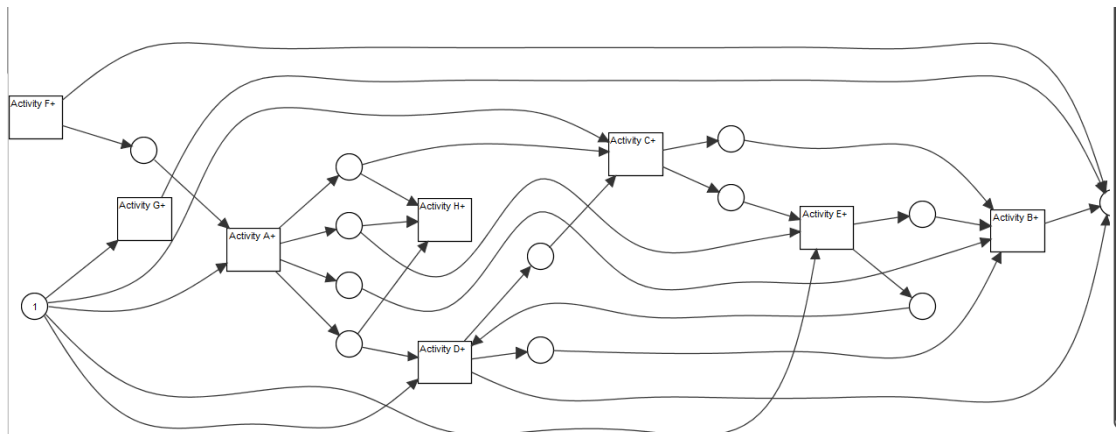
petri net είναι ακόμη πιο πολύπλοκο από αυτό που προέκυψε στο προηγούμενο πείραμα με την χρήση του Inductive Miner κάτι που αποδεικνύει πως οι αύξηση των επαναλήψεων σε συνδυασμό με τον θόρυβο του αντίστοιχου πειράματος επηρέασε αρνητικά την ποιότητα του δικτύου.



Εικόνα 31: Petri Net από το Mine Petri Net With Inductive Miner Plugin για το πείραμα με τις 1000 εκτελέσεις χωρίς θόρυβο

Εκτέλεση Alpha Miner για log με 1000 εκτελέσεις με θόρυβο στα control flows

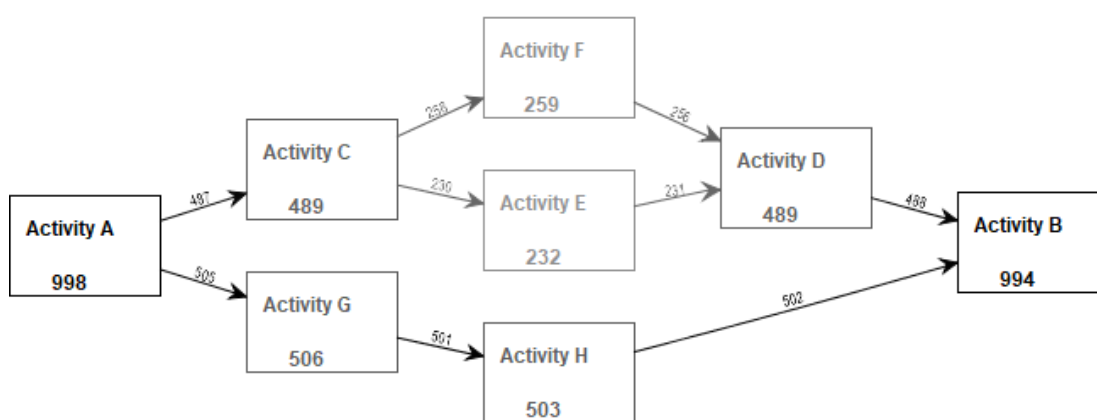
Με την εκτέλεση του Alpha Miner μέσω του Prom plugin “Alpha Miner” είναι αρκετά ευδιάκριτο έχουν αλλάξει οι σχέσεις των δραστηριοτήτων του Petri Net. Στο συγκεκριμένο Petri Net βλέπουμε ότι πέρα από τις αλλαγές στις σχέσεις των δραστηριοτήτων δεν έχουν παρουσιαστεί καθόλου σφάλματα Εικόνες 32.



Εικόνα 32: Petri Net από το Alpha Miner Plugin για το πείραμα με τις 1000 εκτελέσεις με θόρυβο στα control flows

Εκτέλεση Heuristic Miner για log με 1000 εκτελέσεις με θόρυβο στα control flows

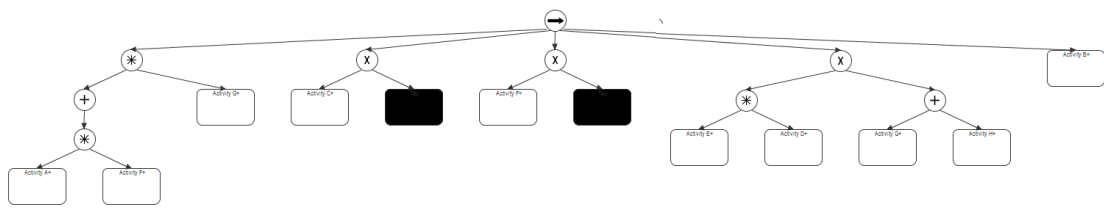
Από το Heuristic Net που προέκυψε σε αυτό το σκέλος του πειράματος φαίνεται πως ο θόρυβος στα control flows του δικτύου δεν επηρέασε την ποιότητα του συγκεκριμένου δικτύου. Φαίνεται πως οι δραστηριότητες σε σειρά έχουν πολύ καλή συχνότητα και πως στις παράλληλες δραστηριότητες οι συχνότητες έχουν διαχωριστεί σχεδόν ισόποσα. Το τρέχον Heuristic Net είναι αρκετά παρόμοιο με το heuristic net από το σκέλος του πειράματος με τις 500 επαναλήψεις χωρίς θόρυβο και θόρυβο μόνο στα control flows αντίστοιχα.



Εικόνα 33: Heuristic Net από το Mine a Heuristic Net with the Heuristic Miner Plugin για το πείραμα με τις 1000 εκτελέσεις με θόρυβο στα control flows

Εκτέλεση Evolutionary tree Miner για log με 1000 εκτελέσεις με θόρυβο στα control flows

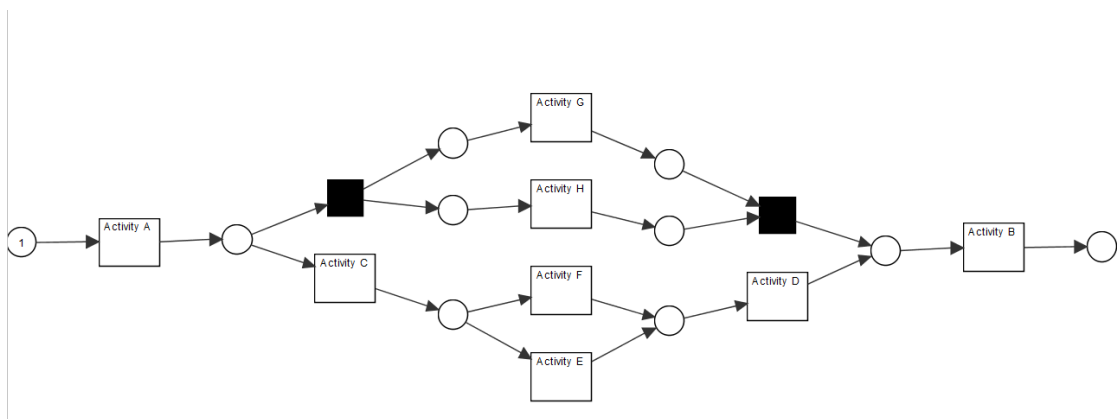
Με το Process Tree που προέκυψε από την εκτέλεση του Evolutionary Tree Miner στο συγκεκριμένο σκέλος του πειράματος με τις 1000 επαναλήψεις παρατηρείται ότι έχουν προκύψει κάποια σφάλματα αλλά και αλλαγές στην δομή του Process Tree σε σχέση με τα προηγούμενα αντίστοιχα πειράματα και τις δραστηριότητες F και G να έχουν χωριστεί σε δύο υπερδραστηριότητες.



Εικόνα 34: Process Tree από το Mine a Process Tree with ETMd Plugin για το πείραμα με τις 1000 εκτελέσεις με θόρυβο στα control flows

Εκτέλεση Inductive Miner για log με 1000 εκτελέσεις με θόρυβο στα control flows

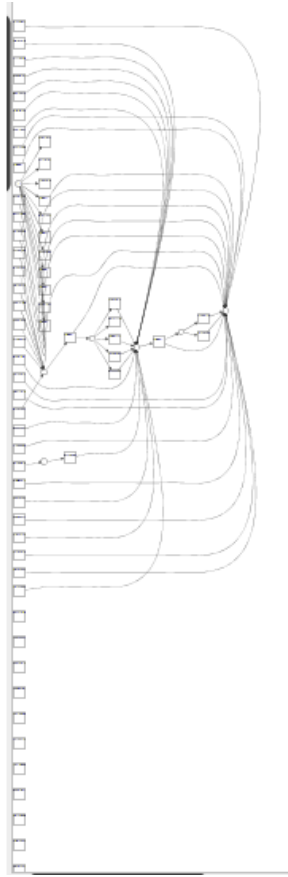
Με την εκτέλεση του Inductive Miner μέσω του αντίστοιχου Prom plugin είναι αρκετά ευδιάκριτο ότι πέρα από την παρουσία δύο σφαλμάτων οι σχέσεις των δραστηριοτήτων του Petri Net δεν έχουν επηρεαστεί.



Εικόνα 35: Petri Net από το Mine a Petri Net with Inductive Miner Plugin για το πείραμα με τις 1000 εκτελέσεις με θόρυβο στα control flows

Εκτέλεση Alpha Miner για log με 1000 εκτελέσεις με πλήρη θόρυβο

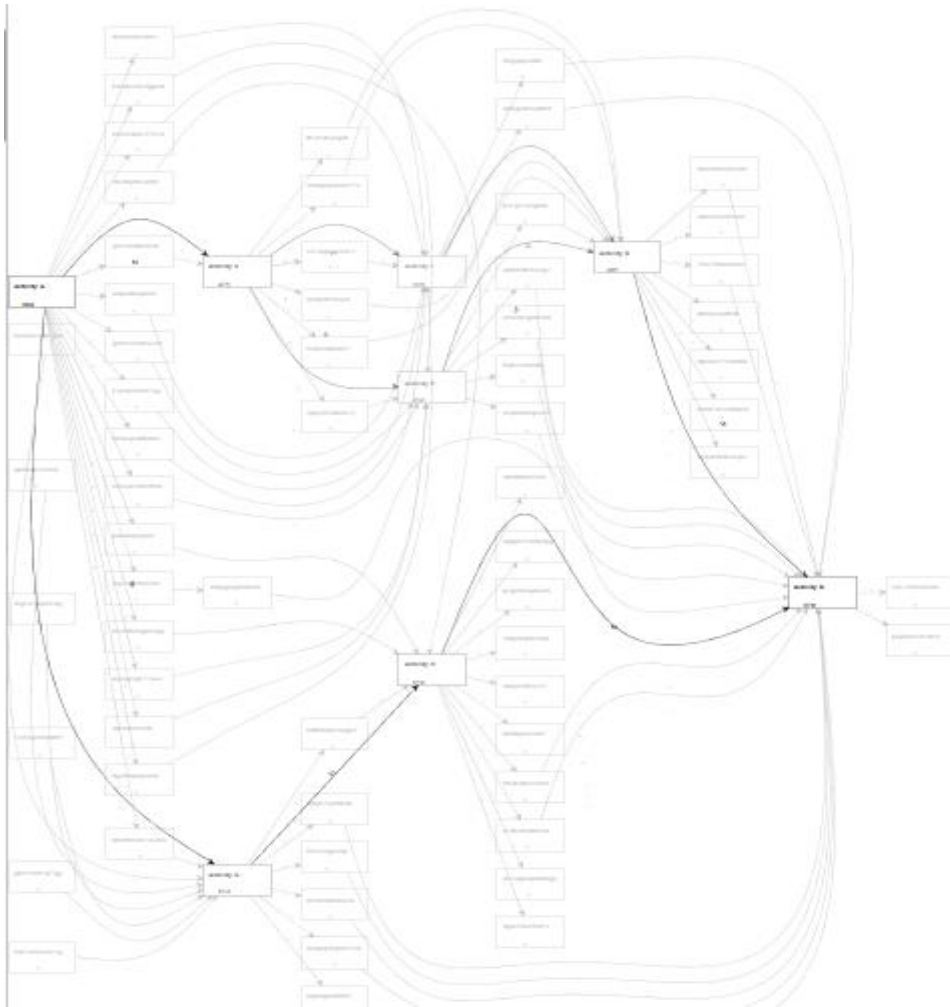
Στο συγκεκριμένο σκέλος του πειράματος έγινε εκτέλεση του Alpha Miner με θόρυβο σε όλο το φάσμα του μοντέλου και προέκυψε ένα petri net με πολλά σφάλματα που το καταστούν εξαιρετικά πολύπλοκο και δυσνόητο και στην συγκεκριμένη περίπτωση παρατηρείται το φαινόμενο μοντέλου σπαγγέτι όπου δεν είναι εμφανείς οι λεπτομέρειες του μοντέλου επίσης για να μπορέσει κάποιος να παρουσιάσει την βασική δομή του μοντέλου θα πρέπει να μεταβεί σε υπερβολικές αφαιρέσεις.



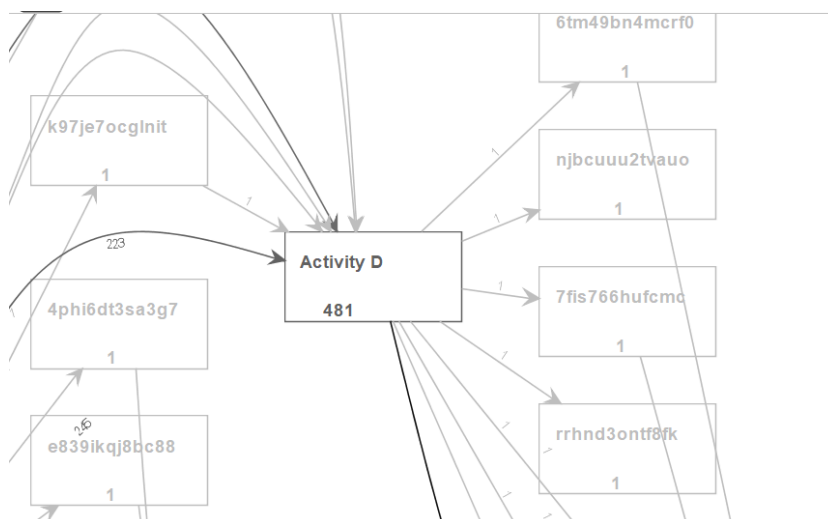
Εικόνα 36: Petri Net από το Alpha Miner Plugin για το πείραμα με τις 1000 εκτελέσεις με πλήρη θόρυβο

Εκτέλεση του Heuristic Miner για log με 1000 εκτελέσεις με πλήρη θόρυβο

Από το Heuristic Net που προέκυψε σε αυτό το σκέλος του πειράματος φαίνεται πως ο πλήρης θόρυβος σε συνδυασμό με τις 1000 εκτελέσεις μείωσε πάρα πολύ την ποιότητα του heuristic net με αποτέλεσμα να είναι αρκετά πολύπλοκο και δυσνόητο. Είναι εμφανής η παρουσία πολλών σφαλμάτων αλλά και η μείωση των συχνοτήτων σε όλες τις δραστηριότητες. Θα μπορούσε κανείς εύκολα να διαπιστώσει το φαινόμενο σπαγγέτι και σε αυτό το δίκτυο. Εικόνα 37



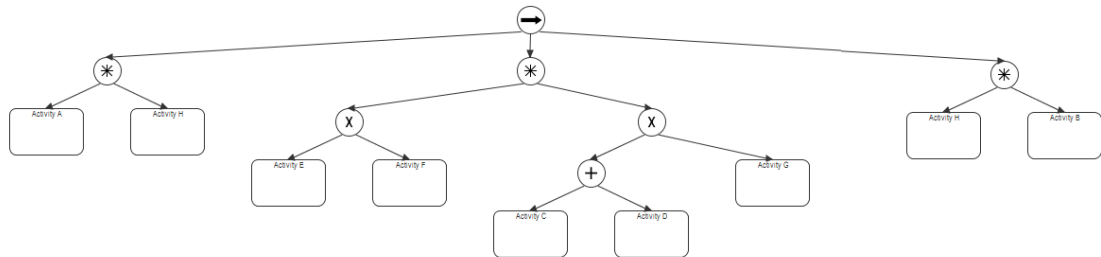
Εικόνα 37: Heuristic Net από το Mine a Heuristic Net with the Heuristic Miner Plugin για το πείραμα με τις 1000 εκτελέσεις με πλήρη θόρυβο



Εικόνα 38: Κομμάτι από το Heuristic Net από το Mine a Heuristic Net with the Heuristic Miner Plugin για το πείραμα με τις 1000 εκτελέσεις με πλήρη θόρυβο

Εκτέλεση του Evolutionary Tree Miner για log με 1000 εκτελέσεις με πλήρη θόρυβο

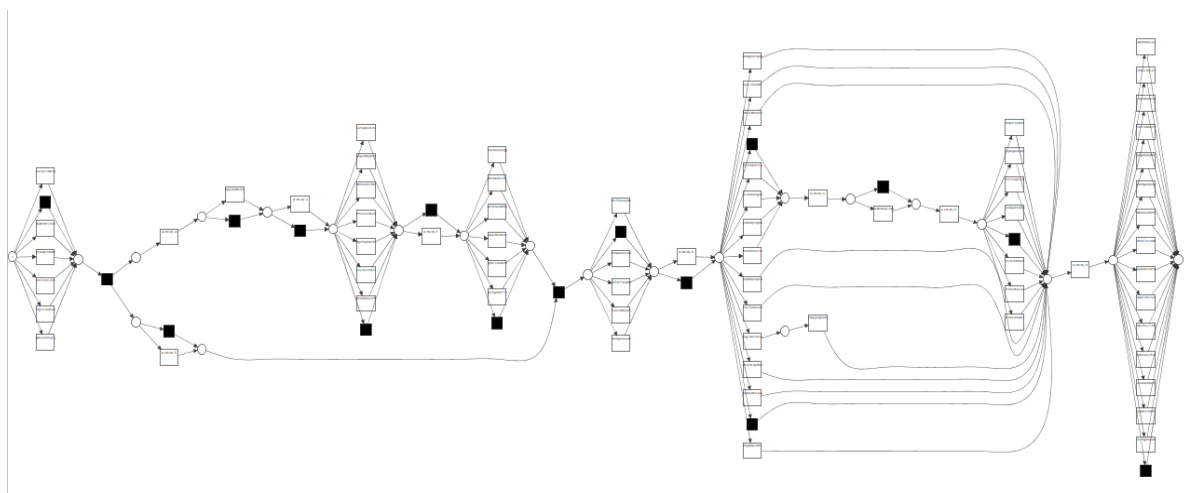
Παρόλο τον θόρυβο και τις 1000 εκτελέσεις φαίνεται πως το δίκτυο που προέκυψε σε αυτό το σκέλος του πειράματος δεν έχει χάσει καθόλου από την ποιότητα του.



Εικόνα 39: Process Tree από το Mine a Process Tree with ETMd Plugin για το πείραμα με τις 1000 εκτελέσεις με πλήρη θόρυβο

Εκτέλεση του Inductive Miner για log με 1000 εκτελέσεις με πλήρη θόρυβο

Με την εκτέλεση του Inductive Miner προκύπτει ένα Petri Net στο οποίο παρουσιάζονται σφάλματα που καταστούν πολύπλοκη την δομή του δικτύου και κατά συνέπεια τις σχέσεις των δραστηριοτήτων. Παρατηρείται ότι το τρέχον petri net είναι ακόμη πιο πολύπλοκο από αυτό που προέκυψε στο προηγούμενο πείραμα με την χρήση του Inductive Miner κάτι που αποδεικνύει πως οι αύξηση των επαναλήψεων σε συνδυασμό με τον θόρυβο του αντίστοιχου πειράματος επηρέασε αρνητικά την ποιότητα και την δομή του δικτύου.



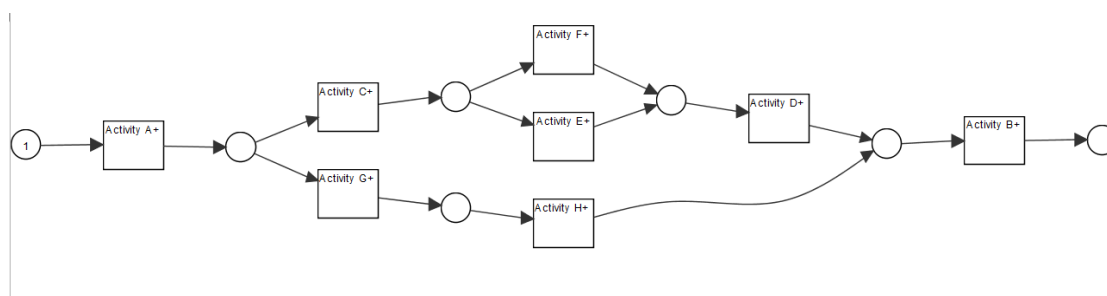
Εικόνα 40: Petri Net από το Mine a Petri Net with Inductive Plugin για το πείραμα με τις 1000 εκτελέσεις με πλήρη θόρυβο

4.5 Πείραμα με 1500 εκτελέσεις

Στο τρέχον πείραμα έγινε εκτέλεση του BPMN μοντέλου για 1500 φορές με την χρήση του λογισμικού PLG2. Με την χρήση του log αρχείου τύπου xes που προέκυψε θα επαναληφθεί η εκτέλεση των τεσσάρων αλγορίθμων με τα αντίστοιχα Prom Plugins.

Εκτέλεση Alpha Miner για log με 1500 εκτελέσεις χωρίς προσθήκη θορύβου

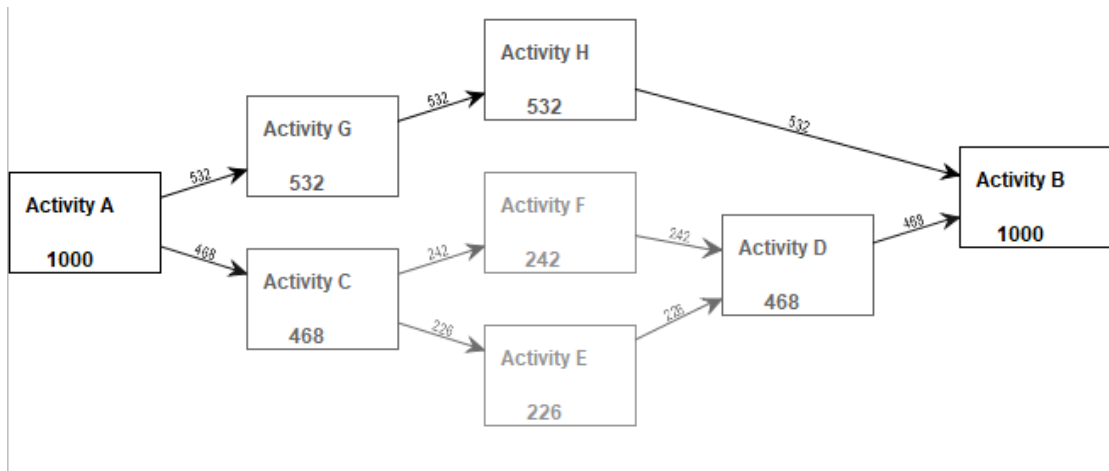
Σε αυτό το σκέλος του πειράματος με την εκτέλεση του Alpha Miner προκύπτει ένα τέλειο ποιοτικά Petri Net στο οποίο δεν παρατηρούνται καθόλου σφάλματα και η εικόνα του θυμίζει το BPMN μοντέλο αλλά και τα αντίστοιχα Petri Net από τα προηγούμενα πειράματα.



Εικόνα 41: Petri Net από το Alpha Miner Plugin για το πείραμα με τις 1500 εκτελέσεις

Εκτέλεση Heuristic Miner για log με 1500 εκτελέσεις χωρίς θόρυβο

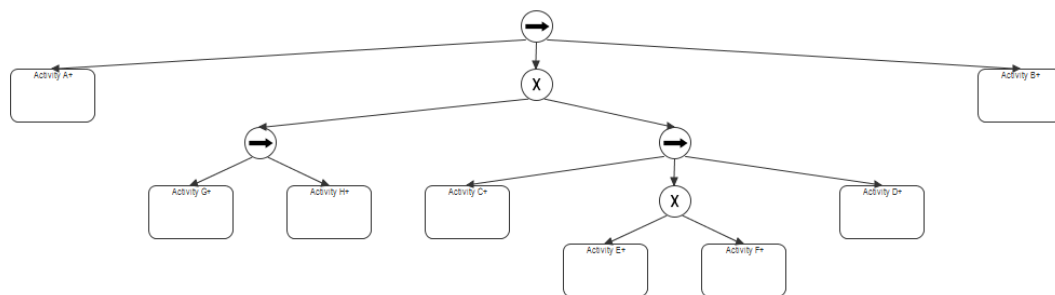
Από την εκτέλεση του Heuristic Miner με την Χρήση του Plugin Mine for a Heuristics Net using Heuristics Miner. Φαίνεται πως οι δραστηριότητες σε σειρά έχουν πολύ καλή συχνότητα και πως στις παράλληλες δραστηριότητες οι συχνότητες έχουν διαχωριστεί σχεδόν ισόποσα. Το τρέχον Heuristic Net είναι αρκετά παρόμοιο με το heuristic net από το σκέλος του πειράματος με τις 500 επαναλήψεις χωρίς θόρυβο και το αντίστοιχο heuristic net από το πείραμα με τις 1000 επαναλήψεις.



Εικόνα 42: Heuristic Net από το Mine a Heuristic Net with the Heuristic Miner Plugin για το πείραμα με τις 1500 εκτελέσεις χωρίς θόρυβο

Εκτέλεση Evolutionary tree Miner για log 1500 εκτελέσεις χωρίς θόρυβο

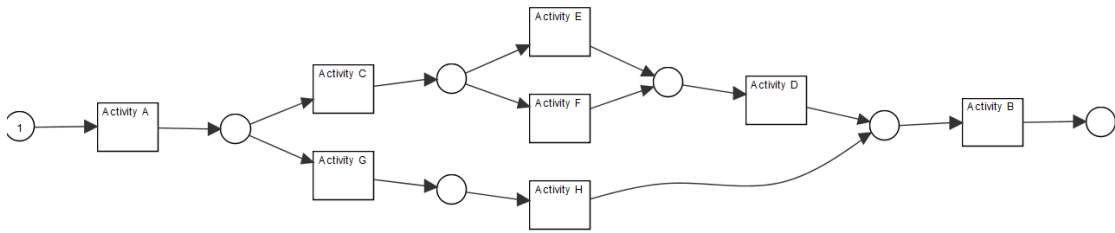
Στο τρέχον σκέλος του πειράματος παράχθηκε ένα Process Tree με την χρήση του Evolutionary Tree Miner μέσω του plugin “Mine a Process Tree with ETMd”. Το τρέχον Process Tree είναι άριστο ποιοτικά και ίδιο με το Process Tree από το σκέλος του πειράματος με τις 500 και 1000 επαναλήψεις χωρίς θόρυβο.



Εικόνα 43: Process Tree από το Mine a Process Tree with ETMd Plugin για το πείραμα με τις 1500 εκτελέσεις και χωρίς θόρυβο

Εκτέλεση Inductive Miner για log με 1500 εκτελέσεις χωρίς θόρυβο

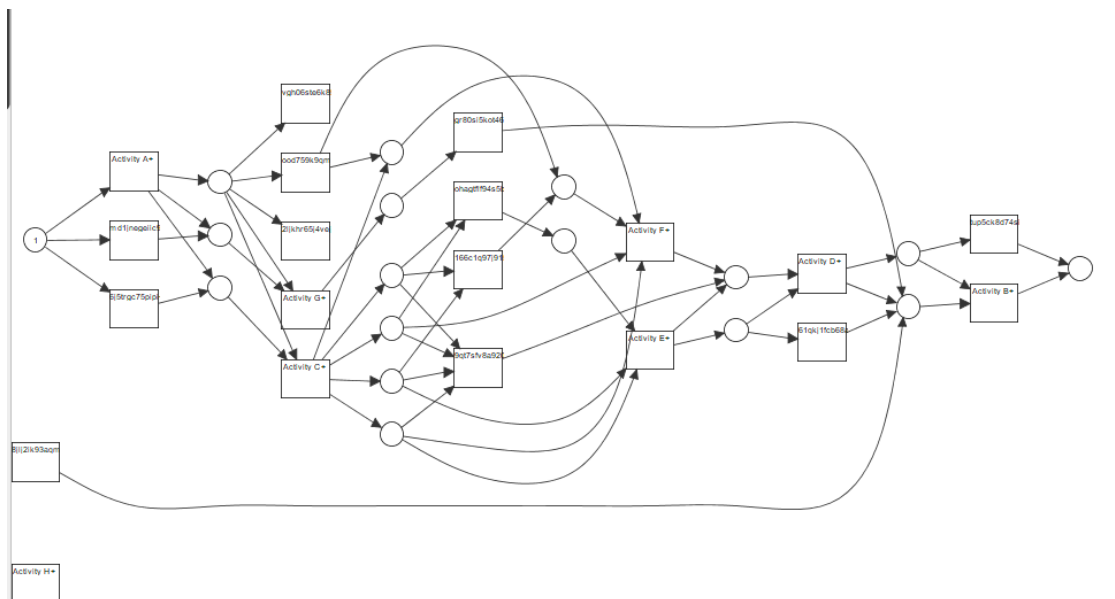
Σε αυτό το σκέλος του πειράματος με την εκτέλεση του Inductive Miner προκύπτει ένα τέλειο ποιοτικά Petri Net στο οποίο δεν παρατηρούνται καθόλου σφάλματα καθώς δεν υπάρχουν μέσα στο event log και η εικόνα του θυμίζει το BPMN μοντέλο αλλά και το αντίστοιχο Petri Net από του προηγούμενου πειράματος.



Εικόνα 44: Petri Net από το Mine Petri Net With Inductive Miner Plugin για το πείραμα με τις 1500 εκτελέσεις χωρίς θόρυβο

Εκτέλεση Alpha Miner για log με 1500 με προσθήκη θορύβου στα activity names

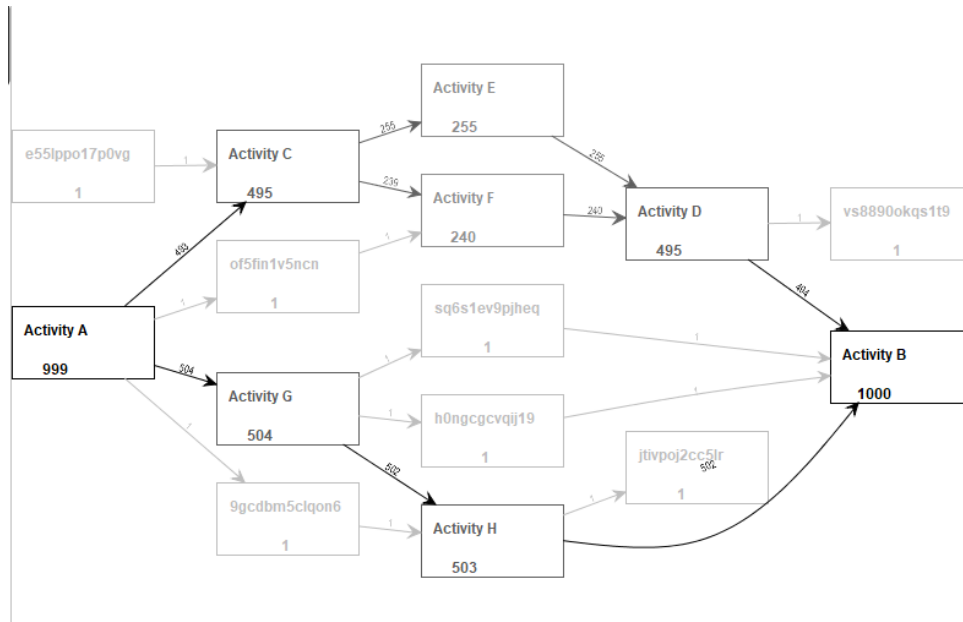
Με την εκτέλεση του Alpha Miner προκύπτει ένα Petri Net στο οποίο παρουσιάζονται σφάλματα που καταστούν περισσότερο πολύπλοκη την δομή του δικτύου και κατά συνέπεια τις σχέσεις των δραστηριοτήτων. Παρατηρείται ότι το τρέχον petri net είναι ακόμη πιο πολύπλοκο από αυτό που προέκυψε στο προηγούμενο πείραμα με την χρήση του Alpha Miner.



Εικόνα 45: Petri Net από το Mine Petri Net With Inductive Miner Plugin για το πείραμα με τις 1500 εκτελέσεις με θόρυβο στα activity names

Εκτέλεση Heuristic Miner για log με 1500 με προσθήκη θορύβου στα activity names

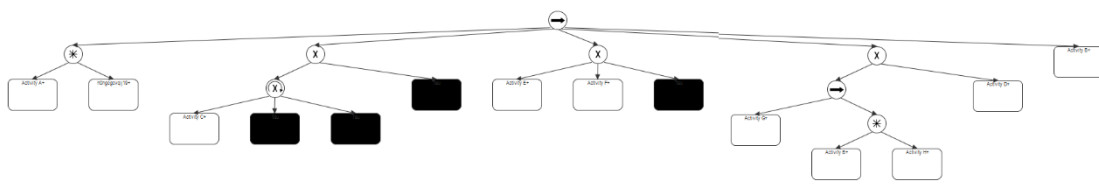
Από το Heuristic Net που προέκυψε σε αυτό το σκέλος του πειράματος φαίνεται πως ο θόρυβος στα activity name σε συνδυασμό με τις 1500 εκτελέσεις μείωσε ακόμη περισσότερο την ποιότητα του heuristic net με αποτέλεσμα να είναι αρκετά πολύπλοκο και δυσνόητο. Είναι εμφανής η παρουσία πολλών σφαλμάτων αλλά και η μείωση των συχνοτήτων σε όλες τις δραστηριότητες.



Εικόνα 46: Heuristic Net από το Mine Heuristics Net with Heuristic Miner Plugin για το πείραμα με τις 1500 εκτελέσεις με θόρυβο στα activity names

Εκτέλεση Evolutionary tree Miner για log με 1500 εκτελέσεις με θόρυβο στα activity names

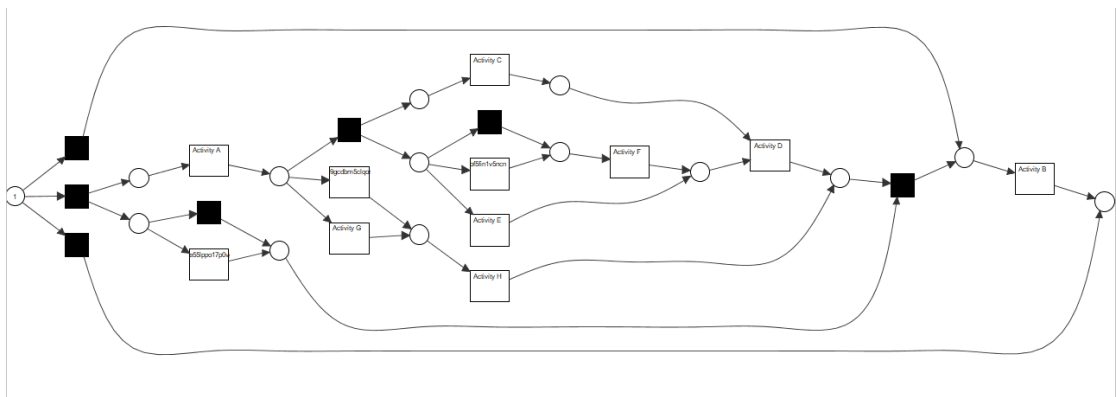
Με το process tree που προκύπτει από την εκτέλεση του Evolutionary tree miner είναι αρκετά προφανές ότι ο θόρυβος στα activity names του μοντέλου σε συνδυασμό με τις 1000 εκτελέσεις προκάλεσαν σφάλματα αλλά επίσης επηρέασε την δομή του process tree.



Εικόνα 47: Process Tree από το Mine a Process Tree with ETMd Plugin για το πείραμα με τις 1000 εκτελέσεις με θόρυβο στα activity names

Εκτέλεση Inductive Miner για log με 1500 εκτελέσεις με θόρυβο στα activity names

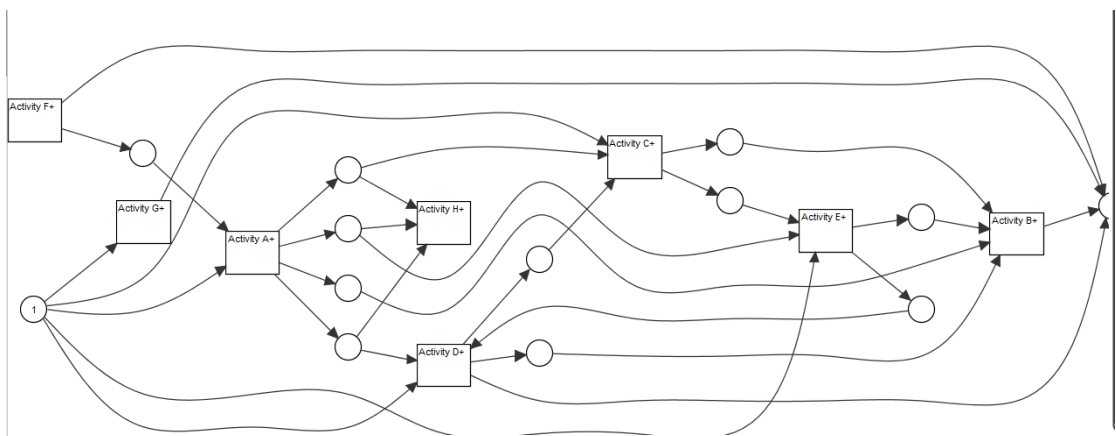
Με την εκτέλεση του Inductive Miner προκύπτει ένα Petri Net στο οποίο παρουσιάζονται σφάλματα που καταστρέφουν πολύπλοκη την δομή του δικτύου και κατά συνέπεια τις σχέσεις των δραστηριοτήτων. Παρατηρείται ότι το τρέχον petri net είναι ακόμη πιο πολύπλοκο από αυτό που προέκυψε στο προηγούμενο πείραμα με την χρήση του Inductive Miner κάτι που αποδεικνύει πως οι αύξηση των επαναλήψεων σε συνδυασμό με τον θόρυβο του αντίστοιχου πειράματος επηρέασε αρνητικά την ποιότητα του δικτύου.



Εικόνα 48: Petri Net από το Mine Petri Net With Inductive Miner Plugin για το πείραμα με τις 1000 εκτελέσεις χωρίς θόρυβο

Εκτέλεση Alpha Miner για log με 1500 εκτελέσεις με θόρυβο στα control flows

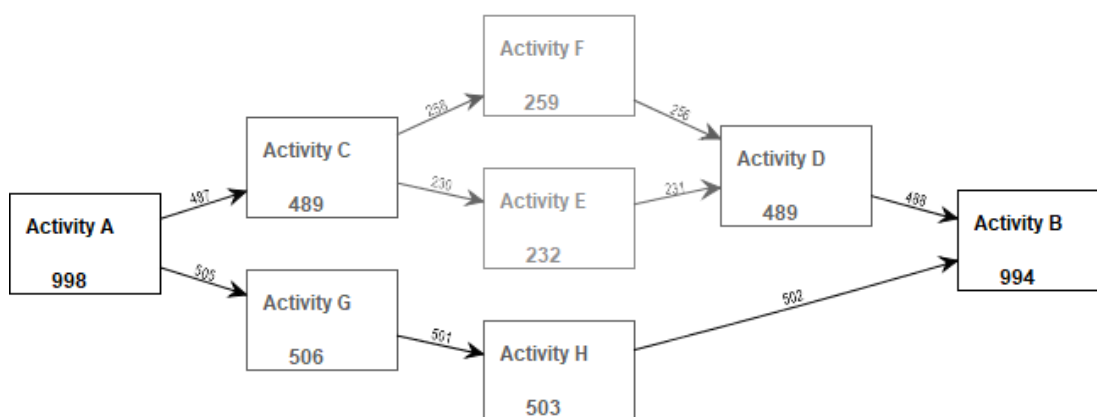
Με την εκτέλεση του Alpha Miner μέσω του Prom plugin “Alpha Miner” είναι αρκετά ευδιάκριτο έχουν αλλάξει οι σχέσεις των δραστηριοτήτων του Petri Net. Στο συγκεκριμένο Petri Net βλέπουμε ότι πέρα από τις αλλαγές στις σχέσεις των δραστηριοτήτων δεν έχουν παρουσιαστεί καθόλου σφάλματα Εικόνες 49.



Εικόνα 49: Petri Net από το Alpha Miner Plugin για το πείραμα με τις 1000 εκτελέσεις με θόρυβο στα control flows

Εκτέλεση Heuristic Miner για log με 1500 εκτελέσεις με θόρυβο στα control flows

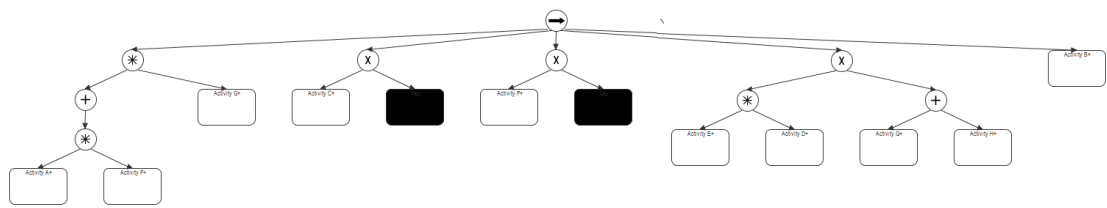
Από το Heuristic Net που προέκυψε σε αυτό το σκέλος του πειράματος φαίνεται πως ο θόρυβος στα control flows του δικτύου δεν επηρέασε την ποιότητα του συγκεκριμένου δικτύου. Φαίνεται πως οι δραστηριότητες σε σειρά έχουν πολύ καλή συχνότητα και πως στις παράλληλες δραστηριότητες οι συχνότητες έχουν διαχωριστεί σχεδόν ισόποσα. Το τρέχον Heuristic Net είναι αρκετά παρόμοιο με το heuristic net από το σκέλος του πειράματος με τις 500 επαναλήψεις χωρίς θόρυβο και θόρυβο μόνο στα control flows αντίστοιχα.



Εικόνα 50: Heuristic Net από το Mine a Heuristic Net with the Heuristic Miner Plugin για το πείραμα με τις 1000 εκτελέσεις με θόρυβο στα control flows

Εκτέλεση Evolutionary tree Miner για log με 1500 εκτελέσεις με θόρυβο στα control flows

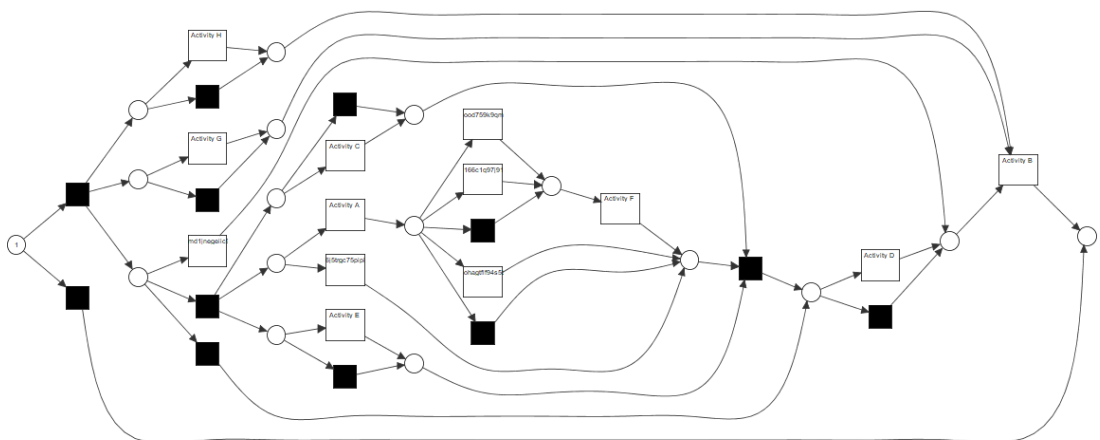
Με το Process Tree που προέκυψε από την εκτέλεση του Evolutionary Tree Miner στο συγκεκριμένο σκέλος του πειράματος με τις 1500 επαναλήψεις παρατηρείται ότι έχουν προκύψει κάποια σφάλματα αλλά και αλλαγές στην δομή του Process Tree σε σχέση με τα προηγούμενα αντίστοιχα πειράματα και τις δραστηριότητες F και G να έχουν χωριστεί δε δύο υπερδραστηριότητες.



Εικόνα 51: Process Tree από το Mine a Process Tree with ETMd Plugin για το πείραμα με τις 1500 εκτελέσεις με θόρυβο στα control flows

Εκτέλεση Inductive Miner για log με 1500 εκτελέσεις με θόρυβο στα control flows

Με την εκτέλεση του Inductive Miner προκύπτει ένα Petri Net στο οποίο παρουσιάζονται σφάλματα που καταστούν πολύπλοκη την δομή του δικτύου και κατά συνέπεια τις σχέσεις των δραστηριοτήτων. Παρατηρείται ότι το τρέχον petri net είναι ακόμη πιο πολύπλοκο από αυτό που προέκυψε στο προηγούμενο πείραμα με την χρήση του Inductive Miner κάτι που αποδεικνύει πως οι αύξηση των επαναλήψεων σε συνδυασμό με τον θόρυβο του αντίστοιχου πειράματος επηρέασε αρνητικά την ποιότητα του δικτύου.



Εικόνα 52: Petri Net από το Mine a Petri Net with Inductive Miner Plugin για το πείραμα με τις 1500 εκτελέσεις με θόρυβο στα control flows

Εκτέλεση Alpha Miner για log με 1500 εκτελέσεις με πλήρη θόρυβο

Στο συγκεκριμένο σκέλος του πειράματος έγινε εκτέλεση του Alpha Miner με θόρυβο σε όλο το φάσμα του μοντέλου και προέκυψε ένα petri net με πολλά σφάλματα που το καταστούν εξαιρετικά πολύπλοκο και δυσνόητο και στην συγκεκριμένη περίπτωση παρατηρείται το φαινόμενο μοντέλου σπαγγέτι στην

χειρότερη μορφή από τις προηγούμενες με τις λεπτομέρειες του μοντέλου επίσης μην είναι εμφανείς.

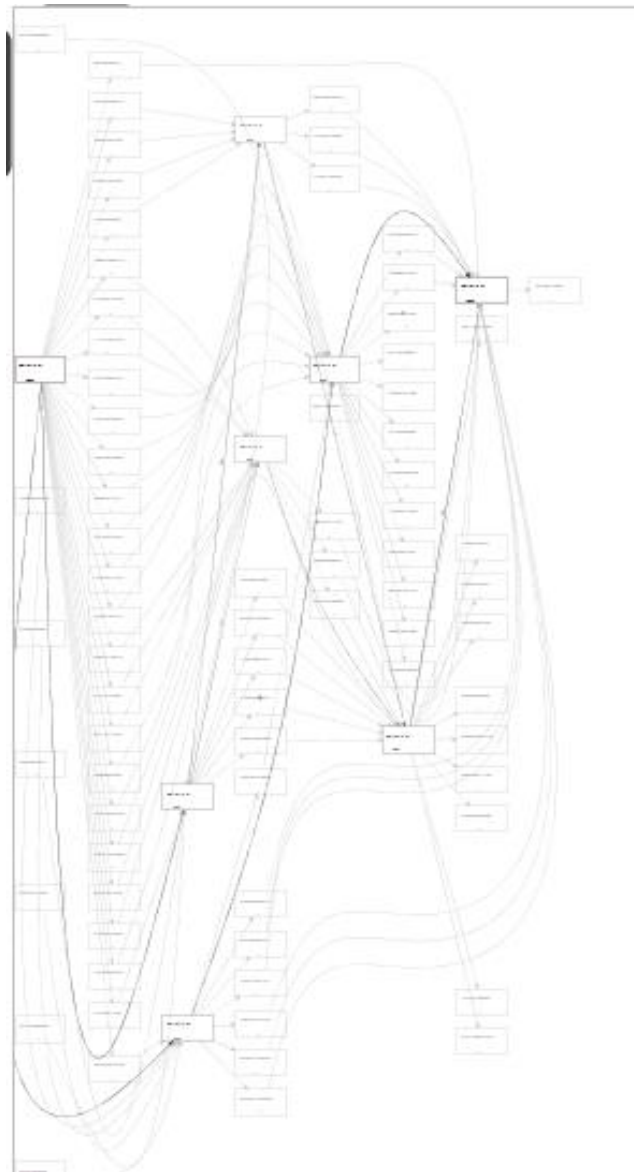


Εικόνα 53: Petri Net από το Alpha Miner Plugin για το πείραμα με τις 1500 εκτελέσεις με πλήρη θόρυβο

Εκτέλεση του Heuristic Miner για log με 1500 εκτελέσεις με πλήρη θόρυβο

Από το Heuristic Net που προέκυψε σε αυτό το σκέλος του πειράματος φαίνεται πως ο πλήρης θόρυβος σε συνδυασμό με τις 1500 εκτελέσεις μείωσε πάρα πολύ την ποιότητα του heuristic net με αποτέλεσμα να είναι αρκετά

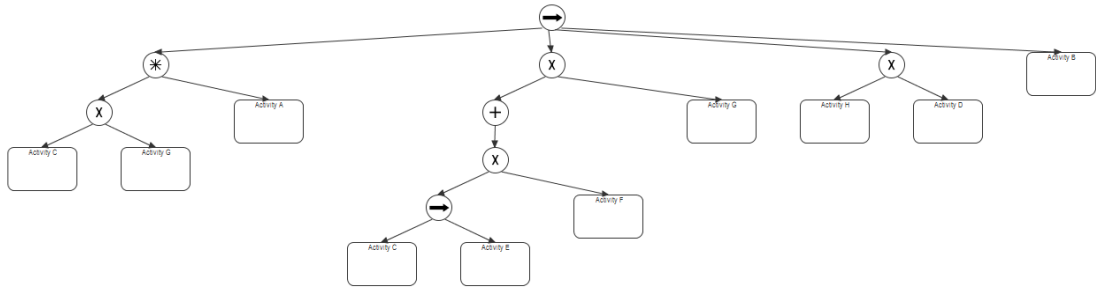
πολύπλοκο και δυσνόητο. Είναι εμφανής η παρουσία πολλών σφαλμάτων αλλά και η μείωση των συχνοτήτων σε όλες τις δραστηριότητες. Θα μπορούσε κανείς εύκολα να διαπιστώσει το φαινόμενο σπαγγέτι και σε αυτό το δίκτυο. Εικόνα 54



Εικόνα 54: Heuristic Net από το Mine a Heuristic Net with the Heuristic Miner Plugin για το πείραμα με τις 1500 εκτελέσεις με πλήρη θόρυβο

Εκτέλεση του Evolutionary Tree Miner για log με 1500 εκτελέσεις με πλήρη θόρυβο

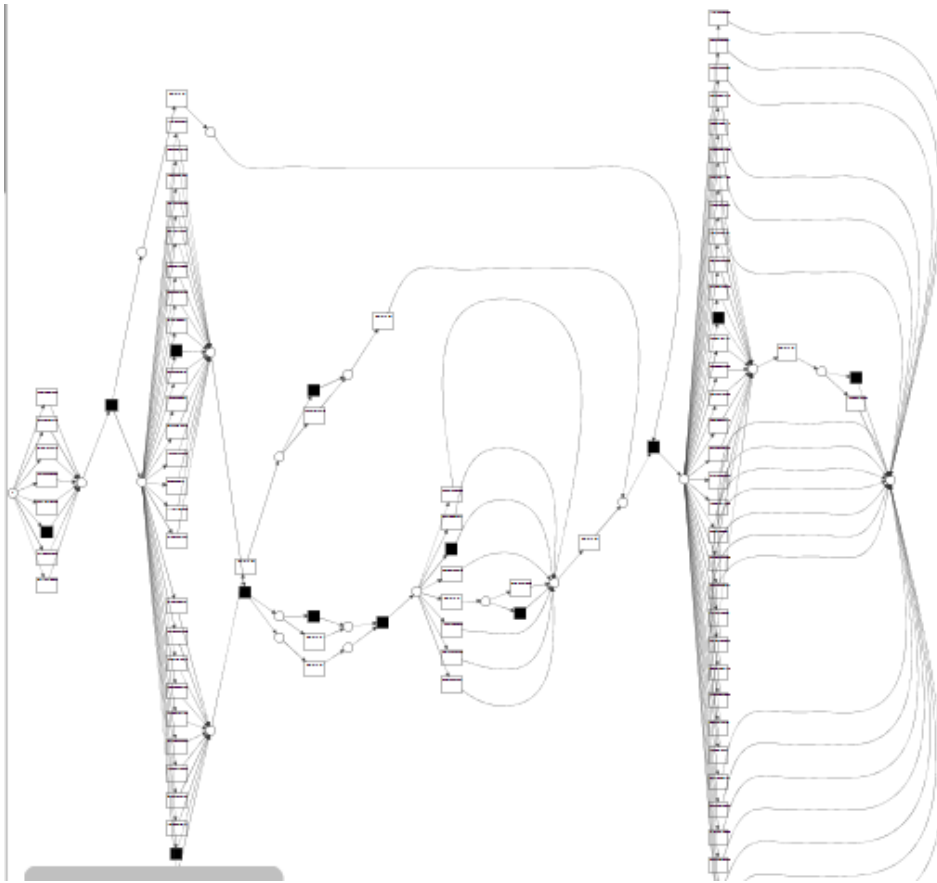
Παρόλο τον θόρυβο και τις 1500 επαναλήψεις φαίνεται πως το δίκτυο που προέκυψε σε αυτό το σκέλος του πειράματος δεν έχει χάσει καθόλου από την ποιότητα του.



Εικόνα 55: Process Tree από το Mine a Process Tree with ETMd Plugin για το πείραμα με τις 1000 εκτελέσεις με πλήρη θόρυβο

Εκτέλεση του Inductive Miner για log με 1500 εκτελέσεις με πλήρη θόρυβο

Με την εκτέλεση του Inductive Miner προκύπτει ένα Petri Net στο οποίο παρουσιάζονται σφάλματα που καταστούν πολύπλοκη την δομή του δικτύου και κατά συνέπεια τις σχέσεις των δραστηριοτήτων. Παρατηρείται ότι το τρέχον petri net είναι ακόμη πιο πολύπλοκο από αυτό που προέκυψε στο προηγούμενο πείραμα με την χρήση του Inductive Miner κάτι που αποδεικνύει πως οι αύξηση των επαναλήψεων σε συνδυασμό με τον θόρυβο του αντίστοιχου πειράματος επηρέασε αρνητικά την ποιότητα του δικτύου επίσης παρατηρείται το φαινόμενο δικτύου σπαγγέτι στην χειρότερη μορφή από όλα τα προηγούμενα δίκτυα.



Εικόνα 56: Petri Net από το Mine a Petri Net with Inductive Miner Plugin για το πείραμα με τις 1500 εκτελέσεις με πλήρη θόρυβο

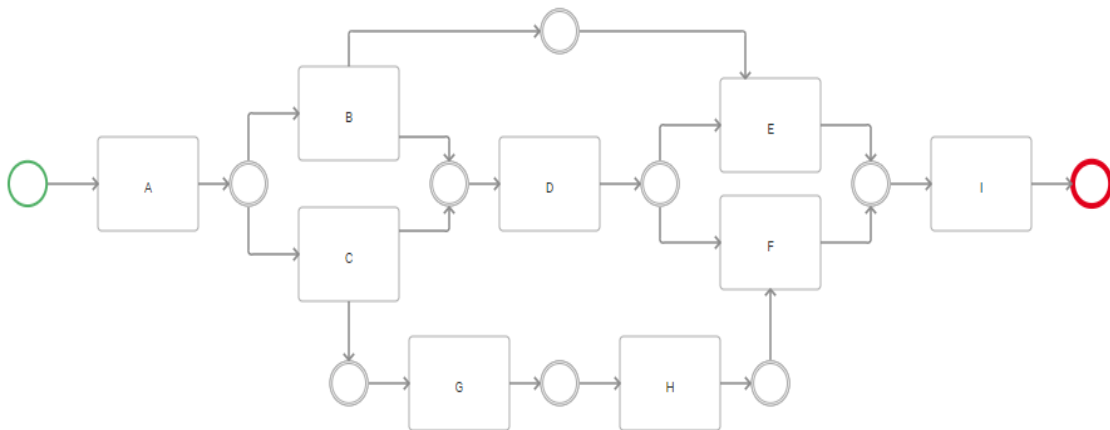
4.6 Evaluation των Process Mining μοντέλων

Παρόλο που έχουν προταθεί πολλές διαφορετικές προσεγγίσεις process mining κατά την τελευταία δεκαετία, δεν υπάρχει διαθέσιμο πρότυπο μέτρο για την αξιολόγηση της ποιότητας ενός τέτοιου μοντέλου μάθησης. Οι μετρήσεις ποιότητας απαιτούνται επειδή ένα μοντέλο δεν μπορεί να εξηγήσει πάντα όλα τα δεδομένα και υπάρχουν πολλαπλά μοντέλα για τα ίδια δεδομένα. Είναι σημαντικό να γίνει κατανοητό ότι δεν υπάρχει ποτέ μόνο ένα ενιαίο μοντέλο για ένα event log. Για να γίνει αυτό λίγο πιο κατανοητό παρακάτω θα παρουσιαστεί ένα παράδειγμα με ένα event log το οποίο θα απεικονιστεί σε τέσσερα διαφορετικά μοντέλα επεξεργασίας με βάση το αρχείο καταγραφής συμβάντων στο αρχικό event log. Πίνακας 3

Event log	
No. of Instances	Log Traces
1207	ABDEI
145	ACDGHFI
56	ACGDHFI
23	ACHDFI
26	ACDHFI

Πίνακας 3 Event Log

Process Model (A)



Εικόνα 57 Process Model (A)

Fitness +
Precision +
Generalization +
Structure +

Στο πρώτο παράδειγμα στην εικόνα 57 παρατηρούμε ότι το μοντέλο διεργασιών μπορεί αναπαράγει όλο το σύνολο των δεδομένων και για αυτό τον λόγο έχει αυξημένο το Fitness. Παρατηρείται επιπλέον ότι δεν υπάρχουν escape edges που έχει ως συνέπεια να έχουμε θετικό πρόσημο και στο precision. Παράλληλα δεν παρατηρούμε δραστηριότητες να μην εκτελούνται συχνά οπότε το generalization έχει και αυτό θετικό πρόσημο όπως και η δομή του (απλότητα).

Process Model (B)

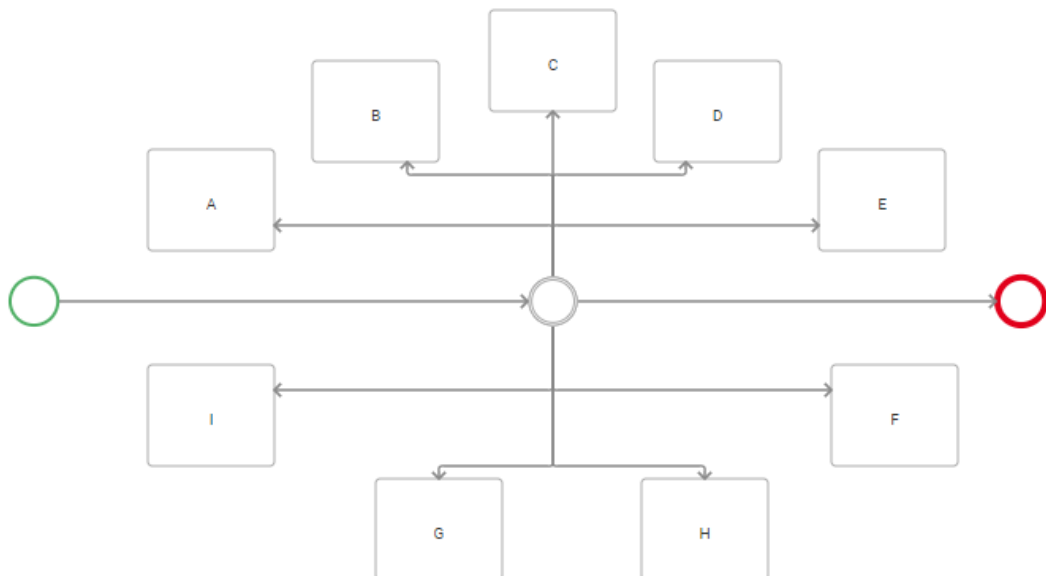


Εικόνα 58: Process Model (B)

Fitness -
Precision +
Generalization -
Structure +

Στο παράδειγμα της εικόνας 58 απεικονίζεται ένα απλό μοντέλο το οποίο δεν έχει την δυνατότητα να παράγει το σύνολο δεδομένων καθώς δεν περιλαμβάνει τα 5 ίχνη. Επιπροσθέτως δεν εμφανίζονται όλες οι δραστηριότητες από το event log.

Process Model (C)



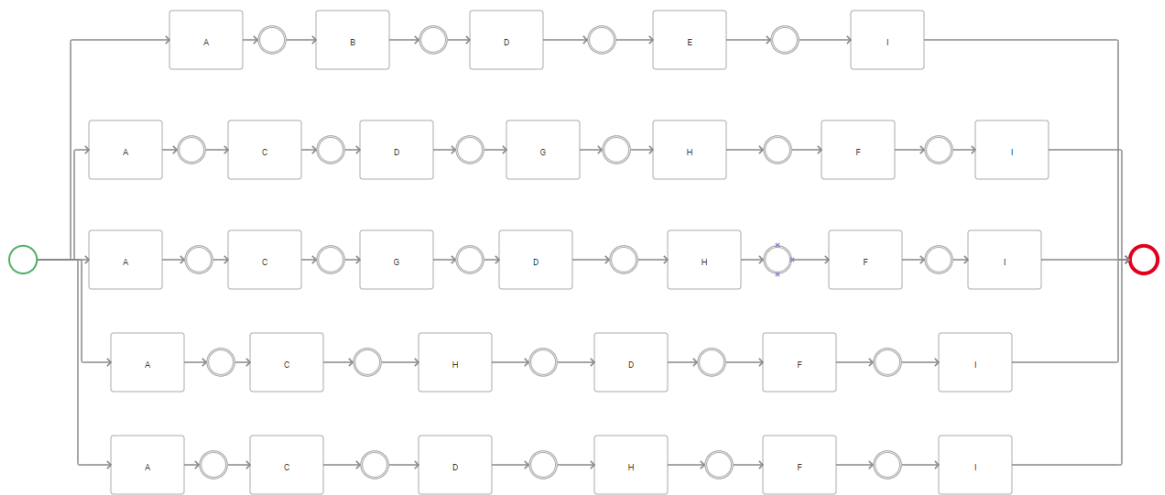
Εικόνα 59: Process Model (C)

Fitness +
Precision -

Generalization +
Structure +

Στο τρίτο παράδειγμα και στην εικόνα 59 παρατηρείται 'ότι το μοντέλο επιτρέπει πολύ περισσότερη συμπεριφορά από ότι στο αρχείο καταγραφής συμβάντων. Αυτό έχει ως αποτέλεσμα όλες οι μετρικές ποιότητας να είναι καλές εκτός το precision. Η ικανότητα του μοντέλου να αναπαριστά όλα τα log traces ακόμα και traces που δεν υπάρχουν στο παράδειγμά μας καθιστούν το μοντέλο ασαφές και μάλλον όχι ικανό να μας εξηγήσει τις διεργασίες που κρύβει.

Process Model (D)



Εικόνα 60: Process Model (D)

Fitness +
Precision +
Generalization -
Structure -

Στο τέταρτο και τελευταίο μοντέλο παρουσιάζονται όλα τα traces σαν γραμμές διεργασιών. Αυτό έχει σαν αποτέλεσμα το fitness και το precision να έχουν θετικά πρόσημα. Από την άλλη μεριά η γενίκευση του μοντέλου καθώς και η εμφάνιση πολλών δραστηριοτήτων πάνω από μια φορά μειώνει την μετρική του generalization καθώς και την δομή.

Παρόλο που, οι αλγόριθμοι προσπαθούν να δημιουργήσουν μοντέλα διεργασιών τα οποία εξισορροπούν πάνω στις μετρικές ποιότητας, έχουν και

κάποια επιπρόσθετα προβλήματα να λύσουν τα οποία παρουσιάζονται παρακάτω.

Προβλήματα που παρουσιάζουν τα μοντέλα διεργασιών.

Η **τέλεια πληρότητα** σε κάποιο event log (να υπάρχουν εκτελέσεις με όλους τους πιθανούς συνδυασμούς δραστηριοτήτων) είναι σχεδόν αδύνατον γιατί το σύνολο των συνδυασμών των παράλληλων δραστηριοτήτων αυξάνεται εκθετικά. Έτσι, είναι γενικά επιθυμητή η γενίκευση πέρα από τις παρατηρούμενες αλληλουχίες για τη διευκόλυνση της ταυτόχρονης ή συνδυασμένης συμπεριφοράς

Η **περαιτέρω αφαίρεση** είναι επίσης ένα πρόβλημα. Παραδείγματος χάριν, παρουσία υπερβολικά πολύπλοκων διαδικασιών, συχνά δεν συντελεί στην εμφάνιση ενός πολύ λεπτομερούς μοντέλου ("σπαγγέτι"). Επιπλέον, κάποιος μπορεί να θέλει να δείξει την κύρια ροή μιας διαδικασίας και έτσι να κάνει πιθανές εξαιρέσεις. Στην περίπτωση αυτή, η αφαίρεση μπορεί να οδηγήσει σε μοντέλα με μειωμένο precision και μειωμένο fitness. Στην πραγματικότητα συχνά υπάρχουν και απλούστερες διαδικασίες (ή τμήματα διαδικασιών) που παρουσιάζουν μόνο διαδοχική δρομολόγηση, εναλλακτική συμπεριφορά και βρόχους (αλλά χωρίς παραλληλισμό). Παραδείγματα τέτοιων διαδικασιών μπορούν να εντοπιστούν στις διοικητικές διαδικασίες που εφαρμόζονται σε δήμους, ασφαλιστικές εταιρείες κ.λπ.

Στην προσέγγισή της παρούσας διπλωματικής εργασία γίνονται λαμβάνονται οι μετρικές ποιότητας fitness, precision, generalization και simplicity και εστιάζουμε σε τέτοιες απλές διαδικασίες. Ελέγχουμε το evaluation των μοντέλων με την υπόθεση ότι τα μοντέλα θα πρέπει να έχουν όσο το δυνατόν καλύτερο fitness (γίνεται αξιολόγηση σχετικά με την ικανότητα του μοντέλου να αναπαραστήσει όλα τα καταγεγραμμένα traces στο event log) και precision (ένα μοντέλο είναι απόλυτα ακριβές αν δεν έχει αναπαράξει παραπάνω δραστηριότητες από αυτές που είναι καταγεγραμμένες στο event log).

4.7 Πίνακες μετρικών ποιότητας

Fitness				
Experiments	AIPHA MINER	Heuristic Miner	Evolutionary Tree Miner	Inductive Miner
500 Traces – Complete Noise	-	0,9969	0	0,79967
500 Traces – Noise only on the Control Flow	0	0,9969	0	0,9613
500 Traces – Noise only on the activity Names	0	0,9991	0,3513	0,8542
500 Traces – No Noise	0,30386	1	0,7868	1
1000 TR - Complete Noise	-	0,9742	0	0,2129
1000 Traces – Noise only on the Control Flow	-	0,9969	0	0,9962
1000 Traces – Noise only on the activity Names	-	0,9991	0	0,7813
1000 Traces – No Noise	-	1	0	1
1500 Traces – Complete Noise	-	0,9799	0	0,206
1500 Traces – Noise only on the Control Flow	-	0,9966	0	0,9955
1500 Traces – Noise only on the activity Names	-	0,9966	0	0,7211
1500 Traces – No Noise	-	1	0	1

Πίνακας 4 Μετρικές Ποιότητας - Fitness

Generalization				
Experiments	AIPHA MINER	Heuristic Miner	Evolutionary Tree Miner	Inductive Miner
500 Traces – Complete Noise	0,9488	0	0,9802	0,7077
500 Traces – Noise only on the Control Flow	0	0	0,91771	0,8318
500 Traces – Noise only on the activity Names	0,3373	0	0,86667	0,6767
500 Traces – No Noise	0	0	0,375	0,2857
1000 TR - Complete Noise	0,84952	0	0,99371	0,94379
1000 Traces – Noise only on the Control Flow	0	0	0,61481	0,94422
1000 Traces – Noise only on the activity Names	-	0	-	0,2857
1000 Traces – No Noise	0,88235	0	0,96415	0,93222
1500 Traces – Complete Noise	0,83319	0	0,98811	0,95255
1500 Traces – Noise only on the Control Flow	0,9488	0	0,9802	0,7077
1500 Traces – Noise only on the activity Names	0	0	0,91771	0,8318
1500 Traces – No Noise	0,3373	0	0,86667	0,6767

Πίνακας 5 Μετρικές Ποιότητας - Generalization

Precision				
Experiments	AIPHA MINER	Heuristic Miner	Evolutionary Tree Miner	Inductive Miner
500 Traces – Complete Noise	0,22706	0	0,7248	0,4267
500 Traces – Noise only on the Control Flow	0	0	0,75	0,8888
500 Traces – Noise only on the activity Names	0,66667	0	0,92381	0,5404
500 Traces – No Noise	0	0	0,96875	1
1000 TR - Complete Noise	0,09118	0	0,47626	0,3327
1000 Traces – Noise only on the Control Flow	0	0	0,28363	0,33349
1000 Traces – Noise only on the activity Names	0,34333	0	0,55769	0,66389
1000 Traces – No Noise	-	0	0,65739	1
1500 Traces – Complete Noise	0,15335	0	0,55879	0,41562
1500 Traces – Noise only on the Control Flow	0	0	0,37446	0,19817
1500 Traces – Noise only on the activity Names	0	0	0,56197	0,19817
1500 Traces – No Noise	0,22706	0	0,7248	0,4267

Πίνακας 6 Μετρικές Ποιότητας - Precision

5 Συμπεράσματα – Μελλοντικές Βελτιώσεις

Στην παρούσα διπλωματική εργασία έγινε εφαρμογή, μελέτη και σύγκριση των διαφόρων αλγορίθμων εξόρυξης δεδομένων. Οι αλγόριθμοι που μελετήθηκαν και θα εφαρμόστηκαν είναι οι εξής : Alpha Miner, Inductive Miner, Evolutionary Tree Miner και Heuristic Miner. Για το προσδιορισμό της ποιότητας των αποτελεσμάτων έγινε καταγραφή και σύγκριση των μετρικών ποιότητας generalization , precision και fitness.

Από τα αποτελέσματα των πειραμάτων προκύπτει το συμπέρασμα πως απαιτείται η απαλλαγή από τον θόρυβο στα πλαίσια των μεγάλων διαδικασιών που χαρακτηρίζονται από ένα μεγάλο αριθμό εκτελέσεων με υψηλό ποσοστό σε σειρά σχέσεων μεταξύ τους. Κατά την εκτέλεση των πειραμάτων με πολλές επαναλήψεις και πλήρη θόρυβο δημιουργήθηκαν μοντέλα διεργασιών τα οποία είχαν την μορφή spaghetti. Το κυριότερο μειονέκτημα των μοντέλων αυτών είναι η δυσκολία άντλησης πληροφοριών. Κατά την διαδικασία προσπάθειας απαλλαγής από τον θόρυβο είναι πολύ πιθανόν να γίνουν περαιτέρω αφαιρέσεις με αποτέλεσμα να παραλειφθούν σημαντικές πληροφορίες και για αυτό τον λόγο παρακάτω αναφέρονται οι πιο ανθεκτικοί αλγόριθμοι όπως καταγράφηκαν από τα πειράματα της παρούσας διπλωματικής.

Προκύπτει πως οι πιο ανθεκτικοί αλγόριθμοι σε σχέση με την ύπαρξη του θορύβου είναι ο Heuristic Miner και ο Evolutionary Tree Miner.

Πιο αναλυτικά στα πειράματα με των Heuristic Miner παρατηρείται πως ο συγκεκριμένος αλγόριθμος ήταν πολύ ανθεκτικός σε σχέση με την μετρική ποιότητας Fitness. Ο συγκεκριμένος αλγόριθμος διατήρησε σχεδόν άριστο Fitness σε όλο το φάσμα των πειραμάτων πάνω από 0,97. Παράλληλα λόγω της μη δυνατότητας από την εφαρμογή του ProM δεν έγινε μέτρηση των μετρικών ποιότητας Precision και generalization για τον συγκεκριμένο αλγόριθμο.

Ο Evolutionary Tree Miner παρατηρήθηκε πως ήταν ο πιο ανθεκτικός αλγόριθμος σε σχέση με τον θόρυβο και την αύξηση των επαναλήψεων σε σχέση με τις μετρικές ποιότητας Precision και Generalization. Ο συγκεκριμένος αλγόριθμος διατήρησε πολύ καλή την μετρική ποιότητας Precision σχεδόν σε όλα τα πειράματα. Από τα αποτελέσματα των πειραμάτων προκύπτει πως κανένα μοντέλο δεν είχε καλά αποτελέσματα σε σχέση με το Generalization

αλλά ο συγκεκριμένος αλγόριθμος είχε τα καλύτερα αποτελέσματα σε σχέση με τους άλλους τρεις αλγόριθμους.

Μελλοντικά θα μπορούσαν να γίνουν αναφορές στους authors των plugins που παρουσίασαν τα σχετικά προβλήματα και να επαναληφθούν τα σκέλη των πειραμάτων που δεν ολοκληρώθηκαν λόγω των προβλημάτων του λογισμικού ProM. Επίσης μελλοντικά στα πλαίσια της συνέχειας των πειραμάτων θα μπορούσαν να γίνουν συγκρίσεις σε αντίστοιχα μοντέλα με ακόμη περισσότερες επαναλήψεις κάτι που ήταν αδύνατον να επιτευχθεί στην παρούσα διπλωματική εργασία λόγω έλλειψης σχετικών πόρων.

Τα υπάρχοντα τεχνικά μέσα για την εξόρυξη διεργασιών όπως και τα εργαλεία όπως το ProM έχουν ωριμάσει ώστε να μπορούν να ανταπεξέρχονται σε προβλήματα όπως είναι οι διεργασίες “spaghetti”. Η ανακάλυψη διεργασιών είναι μάλλον η πιο σημαντική και εμφανής πρόκληση για την εξόρυξη δεδομένων. Όπως έγινε προφανές είναι πολύ δύσκολο να κατασκευαστεί ένα αξιόπιστο μοντέλο διεργασιών από ένα event log που περιέχει θόρυβο ή είναι ημιτελής. Παρόλο που η ευριστική εξόρυξη (Heuristic mining) και η γενετική εξόρυξη (Genetic mining) παρέχουν τεχνικές ανεύρεσης διαδικασιών με περισσότερες δυναμικές, πολλές βελτιώσεις είναι δυνατές για την κατασκευή πιο διαισθητικών μοντέλων που έχουν την δυνατότητα να εξηγήσουν μια συμπεριφορά.

Τέλος, οι νέες προσεγγίσεις εξόρυξης διεργασιών θα πρέπει να επανεξετάσουν τα μοντέλα αναπαράστασης που θα χρησιμοποιούν. Οι περισσότερες προσεγγίσεις χρησιμοποιούν γραφικές παραστάσεις που δίνουν μεγάλη ελευθερία στα μοντέλα τα οποία δεν έχουν πολύ νόημα λόγω των αδιεξόδων και των μη συνδεδεμένων κομματιών. Οι BPMN γραφικές παραστάσεις, μπορούν να εξαλείψουν προβλήματα όπως αυτά σε σχέση με τα Petri Nets. Επιπλέον οι προσεγγίσεις όπως είναι τα δένδρογράμματα και πιο συγκεκριμένα οι αλγόριθμοι που τα χρησιμοποιούν δεν υποφέρουν από τέτοιου είδους προβλήματα. Ωστόσο στα δένδρογράμματα παρατηρούνται δυσκολίες στην έκφραση συγκεκριμένων διεργασιών όπως και στην αποτυχία εμφάνισης διπλότυπων δραστηριοτήτων όπου χρειάζεται. Αυτά τα περιθώρια βελτίωσης εμφανίζουν έναν μεγάλο ερευνητικό χώρο για την ανάπτυξη των μοντέλων εξόρυξης δεδομένων.

6 Βιβλιογραφία

Wil Van Der Aalst (2016) "Process Mining: Data Science in Action" 2nd edition Springer

Carmona, J., van Dongen, B.F., Solti, A., Weidlich, M.(2018) "Conformance Checking Relating Processes and Models" Springer

Burkhart, Thomas; Krumeich, Julian; Werth, Dirk; and Loos, Peter, "Analyzing the Business Model Concept — A Comprehensive Classification of Literature" (2011). *ICIS 2011 Proceedings*.

Stephen A. White, (2004) "Business Process Management Initiative (BPMI)Business Process Modeling Notation (BPMN)" Copyright 2004, BPML.org.

Mark von Rosing, August-Wilhelm Scheer and Henrik von Scheel (2015) "Business Process Model and Notation—BPMN"

Cardoso, Janette, Camargo, Heloisa, (1999) "Fuzziness in Petri Nets" Springer

Andrea Burattin (2016) "PLG2: Multiperspective Process Randomization with Online and Offline Simulations" <http://ceur-ws.org/Vol-1789/bpm-demo-2016-paper1.pdf>

Process Mining Group, Eindhoven University of Technology Process Mining (Definition) [Online] [Cited: October 01, 2019] processmining.org/prom. 24 Aug 2011.

Ana Karla A. de Medeiros, Wil Van der Aalst, A. J. M. M. Weijters, (2003), "Workflow Mining: Current Status and Future Directions" https://www.researchgate.net/publication/221549259_Workflow_Mining_Current_Status_and_Future_Directions

J. M. M. Weijters, Wil Van der Aalst, Ana Karla A. de Medeiros, (2006),
"Process Mining with the Heuristics Miner-algorithm"
https://www.researchgate.net/publication/229124308_Process_Mining_with_the_Heuristics_Miner-algorithm

Wil Van der Aalst, Ana Karla A. de Medeiros, A. J. M. M. Weijters, (2005),
"Genetic Process Mining"
https://www.researchgate.net/publication/220783863_Genetic_Process_Mining

Alejandro Bogarín, Rebeca Cerezo, Cristóbal Romero, (2018) "Discovering learning processes using Inductive Miner: A case study with Learning Management Systems (LMSs)"
https://www.researchgate.net/publication/326743004_Discovering_learning_processes_using_Inductive_Miner_A_case_study_with_Learning_Management_Systems_LMSs

Fabian Rojas Blum, (2015), "Metrics in process discovery"
<https://www.semanticscholar.org/paper/Metrics-in-process-discovery-Blum/5542c60a14129f05baa361a5710524525cb14c17>

Ασπασία Σταμλακού, (2015), "Ανακάλυψη και βελτιστοποίηση επιχειρησιακών διαδικασιών μέσω της μεθόδου του Process Mining"
<http://ikee.lib.auth.gr/record/270228/files/GRI-2015-14856.pdf>

Joos C.A.M. Buijs, Boudewijn F. van Dongen, Wil M.P. van der Aalst, (2012),
"On the Role of Fitness, Precision, Generalization and Simplicity in Process Discovery" https://link.springer.com/chapter/10.1007/978-3-642-33606-5_19

B.F. van Dongen, A.K. Alves de Medeiros, L. Wen, (2009) "Process Mining: Overview and Outlook of Petri Net Discovery Algorithms"
https://link.springer.com/chapter/10.1007%2F978-3-642-00899-3_13

el.wikipedia.org/wiki/Εξόρυξη_διεργασιών

<https://en.wikipedia.org/wiki/Prom>

ΚΟΙΝΩΝΙΑ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ Α.Ε. «Ελληνικό Πλαίσιο Παροχής Υπηρεσιών Ηλεκτρονικής Διακυβέρνησης και Πρότυπα Διαλειτουργικότητας», (2008), “Σημειογραφία των προτύπων BPMN και UML (Activity Diagrams)” Έκδοση 1.00

Sander J.J. Leemans, Dirk Fahland, and Wil M.P. van der Aalst, Exploring, (2015), “Exploring Processes and Deviations”

https://link.springer.com/chapter/10.1007/978-3-319-15895-2_26

Maikel van Eck, BSc, (2013), “Alignment-based Process Model Repair and its Application to the Evolutionary Tree Miner”

<https://pure.tue.nl/ws/portalfiles/portal/46956499/770544-1.pdf>