ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ
ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ

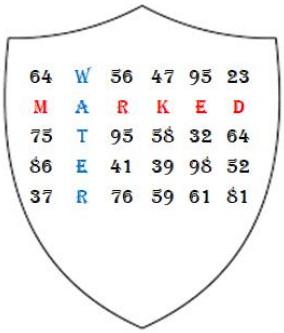
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΑΚΩΝ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΣΤΗΝ ΑΣΦΑΛΕΙΑ ΠΛΗΡΟΦΟΡΙΑΚΩΝ

ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

Προστασία δικαιωμάτων πνευματικής ιδιοκτησίας δεδομένων μέσω τεχνικών υδατοσήμανσης



ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

Αργύρη Παναγιώτη

Επιβλέπουσα : Ακριβή Βλάχου

Μέλη εξεταστικής επιτροπής : Θεόδωρος Κωστούλας, Παναγιώτης Συμεωνίδης

Σάμος, Φεβρουάριος 2023

Η σελίδα αυτή είναι σκόπιμα λευκή.

Πρόλογος και ευχαριστίες

Θέλω να ευχαριστήσω την επιβλέπουσα καθηγήτρια για τις συμβουλές της στην πορεία της εργασίας και την οικογένειά μου για την κατανόηση που έδειξε κατά τη διάρκεια του μεταπτυχιακού. Μέσα από το μεταπτυχιακό γνώρισα τον κόσμο της κυβερνοασφάλειας, ένα κόσμο λογικής και μυστηρίου που συνδυάζει τα μαθηματικά, την πληροφορική τη δημιουργικότητα και την ηθική.

Παναγιώτη Αργύρη
Τμήμα Μηχανικών Πληροφοριακών και Επικοινωνιακών Συστημάτων
ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ

Η σελίδα αυτή είναι σκόπιμα λευκή.

Πίνακας περιεχομένων

Λίστα Σχημάτων

Λίστα Πινάκων

Ακρωνύμια

AES	Advanced Encryption Standard
BAT	Batch file (in the Windows OS)
BCR	Bit correction rate
BER	Bit error rate
BRT	Bit Resetting Techniques
BY	Attribution
©	Copyright
CaaS	Copyright as a service
CC	Creative Commons
CCPA	California Consumer Privacy Act
CIA	Confidentiality-Integrity-Availability
COPE	Create Once Publish Everywhere
CSS	Content Scrambling System
CSV	Comma Separated Values
DB	Database
DIG	Digital Imaging Group
DOI	Digital Objects Identifier
DRM	Digital Rights Management
DSMT	Data Statistics Modifying Techniques
EACML	Extensible Access Control Markup Language
ED	Euclidean Distance
EDM	European Data Model
ERML	Extensible Rights Markup Language
EULA	End User License Agreement
GB	GigaBytes
GDPR	General Data Protection
HD	Hamming Distance
HIPAA	Health Insurance Portability and Accountability Act
IoT	Internet of Things
IPR	Intellectual Property Rights
JSON	JavaScript Object Notation
LIDO	Lightweight Information Describing Objects
LSB	Least Significant Bits

MPEG	Moving Picture Experts Group
MS	Microsoft
MSB	Most Significant Bits
MSE	Mean Squared Error
NC	Non-Commercial
ND	No Derivative Works
ODRL	Open Digital Rights Language
PDF	Portable Document Format
PK	Primary Key
PKB	Primary Key Bits
PSNR	Pean Signal Noise Ratio
RAM	Random Access Memory
RDD	Resilient Distributed Datasets
SA	Share Alike
SaaS	Software as a Service
SP	Service Provider
SQL	Structured Query Language
SSD	Solid State Disk
SSIM	Structural Similarity Index
TRIPS	Trade-Related Aspects of Intellectual Property Rights
UGC	User Generated Content
URL	Uniform Resource Locator
WaaS	Watermarking as a service
WM	Watermark
WRMS	Windows Rights Management Services
XML	Extended Markup Language
ΒΔ	Βάση Δεδομένων
ΓΚΠΔ	Γενικός Κανονισμός Προστασία Δεδομένων της Ευρωπαϊκής Ένωσης
ΠΙ	Πνευματική Ιδιοκτησία
ΠΟΕ	Παγκόσμιος Οργανισμός Εμπορίου

Περίληψη

Η παρούσα διπλωματική εργασία μελετά την προστασία δικαιωμάτων πνευματικής ιδιοκτησίας δεδομένων με νομικά και τεχνολογικά μέσα, συνδυάζοντας βιβλιογραφική έρευνα, υλοποίηση και πειραματική μελέτη ενός δημοφιλούς αλγορίθμου υδατογράφησης δεδομένων.

Η χρήση του διαδικτύου, οι τεχνολογίες Cloud, BigData, IoT και τα ψηφιακά μέσα κοινωνικής δικτύωσης οδήγησαν σε εκθετική αύξηση της ψηφιακής πληροφορίας. Η κοινή χρήση του λογισμικού, των δεδομένων και των ψηφιακών έργων θεωρείται σήμερα δεδομένη, παραβλέποντας πολλές φορές τα δικαιώματα της πνευματικής ιδιοκτησίας. Τα ψηφιακά έργα από τη φύση τους αναπαράγονται, τροποποιούνται και μεταδίδονται εύκολα ευνοώντας την πειρατεία η οποία συνιστά τροχοπέδη στη δημιουργικότητα και στην ανάπτυξη της ψηφιακής οικονομίας.

Η νομοθεσία για την προστασία του δημιουργού και του κατασκευαστή μιας βάσης δεδομένων είναι συγκεκριμένη και αναγνωρίζει ως παραβίαση πνευματικής ιδιοκτησίας τη μη εξουσιοδοτημένη διανομή, αντιγραφή ή οικονομική εκμετάλλευση ψηφιακών έργων. Οι τεχνολογικές λύσεις στην προστασία του δημιουργού συλλογών δεδομένων, αξιοποιούν ιδέες κρυπτογραφίας, στεγανογραφίας και υδατογραφίας που αφορούν την κωδικοποίηση, την απόκρυψη και την ενσωμάτωση πληροφορίας αντίστοιχα. Η υδατογράφηση (watermarking) ειδικότερα, αφορά την τεχνική ενσωμάτωσης ψηφιακών αποτυπωμάτων (marks) για την προστασία των πνευματικών δικαιωμάτων αφού μπορεί να επιβεβαιώσει το δημιουργό, να επαληθεύσει την ακεραιότητα των δεδομένων και να ιχνηλατίσει ένα έργο, ταυτοποιώντας τον πειρατή μέσω του ενσωματωμένου αποτυπώματος στα αντίγραφα.

Το πειραματικό μέρος της εργασίας, υλοποιεί έναν κλασικό αλγόριθμο υδατογράφησης με δυο παραλλαγές του, σχεδιάζει και εκτελεί πολλαπλά σενάρια επιθέσεων σε υδατογραφημένα σύνολα με στόχο να αξιολογήσει την ανθεκτικότητα της προσέγγισης δοκιμάζοντας συνδυασμούς παραμέτρων. Η προσέγγιση υλοποιεί ένα παραμετρικό πρόγραμμα που καλείται επαναληπτικά μέσα από δέσμες ενεργειών (scripts) για την αυτοματοποιημένη εκτέλεση πολλαπλών σεναρίων με διαφοροποιημένες τιμές παραμέτρων. Αυτή η αυτοματοποίηση βοήθησε στην εξαγωγή πολλών γραφημάτων με στόχο την εξαγωγή χρήσιμων συμπερασμάτων.

Η εργασία συνοδεύεται από τον πηγαίο κώδικα Java του παραμετρικού προγράμματος, δέκα εκτελέσιμα σενάρια και γραφήματα από την εκτέλεση των πειραμάτων που έγιναν σε τρία διαφορετικά αριθμητικά σύνολα δεδομένων.

Λέξεις Κλειδιά: *Υδατοσήμανση Δεδομένων, Πνευματικά Δικαιώματα, Υλοποίηση Αλγόριθμου, Αυτοματοποίηση πειραμάτων, Παραμετρικό Λογισμικό*

Abstract

This thesis studies database copyright protection through legal and technological means, combining literature research, implementation and experimental study of a well known watermarking algorithm for data sets.

The use of the Internet, Cloud technologies, BigData, IoT and digital social media have led to an exponential increase in digital information. The sharing of software, data and digital works is now taken for granted, often overlooking intellectual property rights. Digital works by their very nature are easily reproduced, modified and transmitted, favoring piracy which consists an obstacle for creativity and the development of the digital economy.

The legislation to protect the creator and the constructor of a database is very detailed and recognizes the unauthorized distribution, copying or economic exploitation of digital works as an infringement of intellectual property. The technological solutions for the protection of the creator of data collections, utilize ideas of cryptography, steganography and watermarking which concern the coding, concealment and integration of information respectively. Watermarking in particular refers to the technique of embedding digital fingerprints (marks) for the protection of copyright since it can confirm the author, verify the integrity of the data and trace a work, identifying the pirate through the fingerprint embedded in the copies.

The experimental part of the work implements a classical watermarking algorithm with two variations, designs and executes multiple attack scenarios on watermarked sets with the aim of evaluating the robustness of the approach by testing combinations of parameters. The approach implements a parametric program that is executed iteratively through scripts for the automated execution of numerous scenarios with differentiated parameter values. This automation helped to extract many graphs with aim to draw useful conclusions.

The paper is accompanied by the java source code of the parametric program, ten executable attack scenarios and the graphs from the execution of the experiments on three different numerical data sets.

Keywords: *Database Watermarking, Copyright, Algorithm implementation, Experiment Automation, Parametric software*

1 Εισαγωγή

Η ραγδαία ανάπτυξη του διαδικτύου διευκόλυνε την αντιγραφή και τη διανομή ψηφιακού περιεχομένου. Παράλληλα όμως, ευνόησε την παράνομη χρήση και αναπαραγωγή περιεχομένου χωρίς την άδεια του δημιουργού. Τα ψηφιακά δεδομένα από τη φύση τους αναπαράγονται, τροποποιούνται και μεταδίδονται εύκολα, ευνοώντας την πειρατεία η οποία προσβάλλει δικαιώματα πνευματικής ιδιοκτησίας, συνεπώς αποτελεί τροχοπέδη στη δημιουργικότητα και υπονομεύει την ανάπτυξη της οικονομίας των ψηφιακών αγαθών [1]. Η ψηφιακή βιομηχανία της εικόνας, της μουσικής και των βίντεο αντιμετωπίζει εντονότερα τα φαινόμενα της πειρατείας οπότε είναι επιτακτική η ανάγκη για ορθολογική διαχείριση των ψηφιακών δικαιωμάτων με νομικά και τεχνολογικά μέσα. Η ψηφιακή υδατοσήμανση δίνει μια αποτελεσματική τεχνική λύση στην προστασία του ψηφιακού περιεχομένου από μη εξουσιοδοτημένη αναδιανομή και αντιγραφή, αφού αναγνωρίζει τον ιδιοκτήτη. Ταυτόχρονα, λειτουργώντας και ως δακτυλικό αποτύπωμα, ταυτοποιεί τους νόμιμους παραλήπτες του περιεχομένου. Η υδατοσήμανση εισάγει ένα μικρό σήμα ή θόρυβο στο περιεχόμενο με ελεγχόμενο τρόπο ώστε να μην καταστραφεί η αξία του. Το σήμα αυτό είναι δύσκολο να αφαιρεθεί από επιθέσεις χωρίς να καταστραφεί η χρησιμότητα του περιεχομένου ενώ στις περιπτώσεις πολυμεσικού υλικού, η παρέμβαση μετά από ένα όριο γίνεται εύκολα αντιληπτή από έναν παρατηρητικό χρήστη.

Η πρόσβαση στις ψηφιακές πληροφορίες μέσω της δημιουργίας αντιγράφων αποτελεί μια ιδιαιτερότητα των ψηφιακών μέσων έναντι των παραδοσιακών. Για παράδειγμα, η ανάγνωση ενός αρχείου από μια πηγή στο διαδίκτυο απαιτεί την αναπαραγωγή του περιεχομένου σε προσωρινή τοποθεσία στο σκληρό δίσκο, στη μνήμη του Η/Υ και στην οθόνη όπου προβάλλεται [1]. Η πρόσβαση λοιπόν είναι ισχυρά συνυφασμένη με την αντιγραφή και αποτελεί το αντικείμενο της προστασίας της πνευματικής ιδιοκτησίας. Το δικαίωμα της πρόσβασης ελέγχεται ξεχωριστά από το δικαίωμα της αναπαραγωγής στα δημιουργήματα με υλική υπόσταση. Αντίθετα, η πρόσβαση στις ψηφιακές δημιουργίες δεν μπορεί να αποτρέψει την αντιγραφή. Ιδιαίτερη πρόκληση αποτελεί η παροχή ελεύθερης πρόσβασης με ταυτόχρονο έλεγχο της αναπαραγωγής εστιασμένο στην προστασία της πνευματικής ιδιοκτησίας. Η ψηφιακή αναπαραγωγή με τα σημερινά σύγχρονα μέσα επικοινωνίας, μεταφοράς και αποθήκευσης είναι εύκολη και οικονομική αφού ένα ψηφιακό οπτικοακουστικό υλικό video 2-4 ωρών, μεγέθους 4-8GB, μεταφέρεται μέσα σε λίγα λεπτά στο διαδίκτυο. Η χρήση του διαδικτύου οδηγεί σε εκθετική αύξηση του όγκου της πληροφορίας που

μπορεί εύκολα να διαχυθεί και να αναπαραχθεί μέσα από τις τεχνολογίες cloud, τα μέσα κοινωνικής δικτύωσης και τα δίκτυα. Η κοινή χρήση και διάδοση της πληροφορίας μέσω του διαδικτύου, οι τεχνολογίες cloud και τα ηλεκτρονικά μέσα ηλεκτρονικής δικτύωσης είναι δεδομένα στην εποχή μας. Καθώς οι παραβιάσεις πνευματικής ιδιοκτησίας περιλαμβάνουν τη μη εξουσιοδοτημένη διανομή, την αντιγραφή και την πώληση, η ταυτοποίηση ενός πειρατή είναι σημαντική για την τιμωρία του ενόχου ή την απαλλαγή ενός αθώου που κατηγορείται.

Η προστασία των πνευματικών δικαιωμάτων αξιοποιεί τις τεχνικές της απόκρυψης, της κωδικοποίησης και της ενσωμάτωση κρυφών μηνυμάτων μέσα σε περιεχόμενο. Οι τρεις αυτές τεχνικές ξεχωριστά αλλά και συνδυαστικά, χρησιμοποιούνται στο σύγχρονο ψηφιακό κόσμο για την επίτευξη της ασφάλειας, της προστασίας και της ιδιωτικότητας των ψηφιακών αγαθών. Κοινούς σκοπούς εξυπηρετούν η κρυπτογραφία, η στεγανογραφία και η υδατογραφία, μέσα από διαφορετικές προσεγγίσεις όπως δείχνει και η Εικόνα 1, ενώ το παράρτημα ΙΙ περιλαμβάνει περισσότερα ιστορικά στοιχεία για αυτές.

Εικόνα 1 : Κρυπτογραφία, στεγανογραφία και υδατογραφία, λύσεις εμπιστευτικότητας

1.1 Ο σκοπός της διπλωματικής

Σκοπός της διπλωματικής είναι η επισκόπηση των νομικών και των τεχνολογικών μέτρων για την προστασία της πνευματικής ιδιοκτησίας και των πνευματικών δικαιωμάτων που αφορούν συλλογές δεδομένων και ψηφιακά σύνολα. Η εργασία επικεντρώνεται στα τεχνολογικά μέσα, περιλαμβάνει βιβλιογραφική επισκόπηση της τελευταίας εικοσαετίας, ενώ το πειραματικό μέρος περιλαμβάνει την υλοποίηση ενός γνωστού αλγόριθμου υδατογράφησης [21] με δυο παραλλαγές του για αριθμητικά σύνολα δεδομένων ακεραίων και δεκαδικών αριθμών και την εκτέλεση μεγάλου πλήθους επιθέσεων σε τρία σύνολα δεδομένων. Στο πλαίσιο αυτό :

i. Εξετάζεται ο βαθμός δυσκολίας στην υλοποίηση υδατοσήμανσης σε αριθμητικά δεδομένα.
ii. Αξιολογείται η ανθεκτικότητα της υδατοσήμανσης μέσα από σενάρια και πειράματα που περιλαμβάνουν επιθέσεις με διαφοροποιημένες παραμέτρους.

1.2 Η μεθοδολογία και τα παραδοτέα της διπλωματικής

Η εργασία περιλαμβάνει θεωρητικό και πρακτικό μέρος. Το θεωρητικό μέρος εξερευνά βιβλιογραφικές πηγές από δύο περιοχές, το χώρο της πνευματικής ιδιοκτησίας και την έρευνα στις τεχνικές υδατοσήμανσης συνόλων δεδομένων. Το πρακτικό μέρος σχεδιάζει και υλοποιεί υδατογραφήσεις και σενάρια επιθέσεων αξιολογώντας στη συνέχεια την αποτελεσματικότητα των ανιχνεύσεων. Τα παραδοτέα που συνοδεύουν την εργασία περιλαμβάνουν:

i. Τον πηγαίο κώδικα Java και το εκτελέσιμο του παραμετρικού προγράμματος που εκτελεί υδατογραφήσεις, επιθέσεις και ανιχνεύσεις
ii. Τεκμηρίωση και ελέγχους λειτουργίας του λογισμικού.
iii. Δέκα εκτελέσιμα σενάρια μορφής MS Window scripts (αρχεία με επέκταση .bat).
iv. Τα τρία αριθμητικά σύνολα δεδομένων που χρησιμοποιήθηκαν στα πειράματα.
v. Τα αρχεία καταγραφής και τα γραφήματα των πειραμάτων.

Η γραφική απεικόνιση των αποτελεσμάτων βοήθησε σε μεγάλο βαθμό στον έλεγχο του παραμετρικού προγράμματος, των σεναρίων και στην εξαγωγή χρήσιμων συμπερασμάτων.

1.3 Τα ερευνητικά ερωτήματα

Τα ερευνητικά ερωτήματα στα οποία στοχεύει να απαντήσει η εργασία ακολουθούν.

1. Ποια είναι τα γνωστά νομικά και τεχνολογικά μέτρα προστασίας της πνευματικής ιδιοκτησίας συλλογών δεδομένων ;
2. Υπάρχει τρόπος αναγνώρισης της παράνομης διακίνησης αντιγράφων συλλογών;
3. Ποιες είναι οι κατάλληλες παράμετροι για τη δημιουργία ενός υδατογραφήματος ανθεκτικού σε διάφορες μορφές επιθέσεων ;

1.4 Η δομή της διπλωματικής

Η διπλωματική περιλαμβάνει επτά κεφάλαια, τις βιβλιογραφικές πηγές και τρία παραρτήματα. Το παρόν, Κεφάλαιο 1 αποτελεί μια εισαγωγή. Το Κεφάλαιο 2 παρουσιάζει εργασίες, άρθρα και δημοσιεύσεις των τελευταίων δυο δεκαετιών, σχετικές με την προστασία της πνευματικής ιδιοκτησίας και τις τεχνικές υδατοσήμανσης συνόλων δεδομένων. Το Κεφάλαιο 3 αναλύει την έννοια της πνευματικής ιδιοκτησίας και τις γνωστές προσεγγίσεις προστασίας των πνευματικών δικαιωμάτων του δημιουργού και του κατασκευαστή μιας βάσης δεδομένων. Το Κεφάλαιο 4 ταξινομεί και αναλύει τις τεχνικές υδατοσήμανσης συνόλων δεδομένων, τις επιθέσεις και την ανθεκτικότητά τους σε αυτές. Το Κεφάλαιο 5 εξηγεί τη σχεδίαση και υλοποίηση ενός παραμετρικού λογισμικού υδατογράφησης και επιθέσεων, τα αυτοματοποιημένα σενάρια επιθέσεων και δοκιμών ανθεκτικότητας στα επιλεγμένα σύνολα δεδομένων. Το Κεφάλαιο 6 επεξηγεί τα πειράματα που εκπονήθηκαν και μετρά την ανθεκτικότητα του επιλεγμένου σχήματος υδατογράφησης μέσα από σενάρια επιθέσεων. Το Κεφάλαιο 7 απαντά στα ερευνητικά ερωτήματα της εργασίας, παραθέτει συμπεράσματα και προτάσεις για μελλοντικές εφαρμογές.

Το παράρτημα Ι παραθέτει τα γραφήματα των πειραμάτων ως εικονίδια. Το παράρτημα ΙΙ περιλαμβάνει τεκμήρια από την εκπόνηση της εργασίας. Το παράρτημα ΙΙΙ δίνει μια ιστορική αναφορά της κρυπτογραφίας, της στεγανογραφίας και της υδατογραφίας.

2 Βιβλιογραφική Επισκόπηση

Η βιβλιογραφική επισκόπηση καλύπτει τα νομικά και τα τεχνολογικά μέσα προστασίας της πνευματικής ιδιοκτησίας δεδομένων ενώ επικεντρώνεται στην τεχνική προσέγγιση της υδατογράφησης σε αριθμητικά και επιστημονικά δεδομένα. Η βιβλιογραφική έρευνα αξιοποίησε πρωτεύουσες πηγές καθώς και δευτερεύουσες πηγές των σημαντικών άρθρων ενώ η αναζήτηση έγινε μέσω του ηλεκτρονικού συστήματος αναζήτησης «Google Scholar».

2.1 Σχετικά με την πνευματική ιδιοκτησία

Το ηλεκτρονικό σύγγραμμα του Δ.Τσώλη [1] καλύπτει ένα μεγάλο μέρος των θεμάτων προστασίας της πνευματικής ιδιοκτησίας ψηφιακών δεδομένων.

Η οδηγία 96/9/ΕΟΚ [2,51] είναι μια πολύ βασική νομοθετική πρωτοβουλία για τη νομική προστασία κάθε μορφής βάσεων δεδομένων που αναγνωρίζεται ως πνευματικό δημιούργημα και προστατεύεται από τη νομοθεσία της πνευματικής ιδιοκτησίας. Σύμφωνα με την οδηγία, δημιουργός μιας βάσης δεδομένων θεωρείται το φυσικό πρόσωπο ή την ομάδα που έχει δημιουργήσει τη βάση δεδομένων ή το νομικό πρόσωπο που ορίζεται ως δικαιούχος. Στην ίδια πηγή [2] αναφέρεται ότι ο δημιουργός της βάσης δεδομένων έχει το αποκλειστικό δικαίωμα να εκτελεί ή να επιτρέπει :

i. την προσωρινή ή διαρκή αναπαραγωγή του συνόλου της βάσης δεδομένων ή μέρους της, με οιοδήποτε μέσο ή μορφή
ii. τη μετάφραση, προσαρμογή, διευθέτηση και οποιαδήποτε μετατροπή της βάσης δεδομένων
iii. τη διανομή της βάσης δεδομένων σε οποιαδήποτε μορφή, ή αντιγράφων της
iv. την ανακοίνωση, επίδειξη, αναπαραγωγή ή παρουσίαση της βάσης δεδομένων

Επιπρόσθετα, η ένωση Ευρωπαϊκών ψηφιακών δικαιωμάτων [3], που αποτελείται από ένα δίκτυο 36 οργανισμών πολιτών και προστασίας των ανθρώπινων δικαιωμάτων σε 21 κράτη μέλη της ΕΕ, στοχεύει στην προώθηση και προστασία των δικαιωμάτων και των ελευθεριών στο ψηφιακό περιβάλλον. Οι δέκα αρχές που διέπουν τη λειτουργία της, αποτυπώνονται σε ένα χάρτη ψηφιακών δικαιωμάτων. Κάποιες αρχές αφορούν τα πνευματικά δικαιώματα στην εξάπλωση της ανοιχτής γνώσης. Πιο συγκεκριμένα, οι αρχές [3] αφορούν στην προώθηση της διαφάνειας τη πρόσβασης στην πληροφορία και τη συμμετοχή των πολιτών, στην προστασία των δεδομένων και

της ιδιωτικότητας, στην προσβασιμότητα σε υπηρεσίες του διαδικτύου, στο σεβασμό των πνευματικών δικαιωμάτων και στην προώθηση της ανοιχτής γνώσης που δεν υπόκειται σε δικαιώματα δημιουργών, στην αντίσταση σε μέτρα παρακολούθησης που ξεπερνούν το μέτρο της αναλογικότητας, στην ανωνυμία και ενδυνάμωση της ιδιωτικότητας, στην αντίθεση σε ιδιωτικά μέτρα εξαναγκασμού, στην προώθηση του ανοιχτού λογισμικού, στην υποστήριξη της δημοκρατίας και του κράτους δικαίου και τέλος στον πλουραλισμό και στην ελεύθερη ενημέρωση.

Η πηγή [6] αναφέρεται στα πνευματικά δικαιώματα δημιουργών ψηφιακών έργων, ενώ η πηγή [7], περιγράφει τη δημιουργία και την εξέλιξη των εθνικών στρατηγικών κυβερνοασφάλειας στην προστασία των πνευματικών δικαιωμάτων και της ιδιωτικότητας. Συνοψίζει τις στρατηγικές και τα νομοθετήματα των G34 ισχυρότερων οικονομικά χωρών του κόσμου και αναφέρει ότι το 32% αυτών ασχολείται επίσημα με την προστασία της πνευματικής ιδιοκτησίας. Αντίστοιχα η πηγή [8] αναλύει τα δικαιώματα πνευματικής ιδιοκτησίας της πληροφορίας και συγκρίνει την προστασία των βάσεων δεδομένων στην Ευρώπη και στην Αμερική. Η πηγή [9] αναλύει τη συγκρουσιακή σχέση ανάμεσα στην ανοιχτή πρόσβαση και στην πνευματική ιδιοκτησία των πρωτότυπων έργων που στηρίζει τη δημιουργικότητα και τη διάδοση της γνώσης.

Η πηγή [10] αναλύει και διακρίνει τις έννοιες της *πνευματικής ιδιοκτησίας*, της κλειστής και της ανοιχτής *αδειοδότησης*. Η πνευματική ιδιοκτησία ανήκει στο δημιουργό χωρίς να απαιτείται κάποια ειδική ενέργεια κατοχύρωσης ενώ περιλαμβάνει το ηθικό δικαίωμα της αναφοράς στο δημιουργό καθώς και τα περιουσιακά δικαιώματα της αναπαραγωγής, της επεξεργασίας της διάθεση και της οικονομικής εκμετάλλευσης των έργων του. Η αδειοδότηση από την άλλη, αποτελεί την άσκηση του δικαιώματος εκμετάλλευσης της πνευματικής ιδιοκτησίας (Ν. 2121/1993) και απαιτεί μια σειρά από ενέργειες.

Υπάρχουν δυο τύποι αδειών, οι κλειστές και οι ανοιχτές. Στις κλειστές άδειες ή εκχωρήσεις δικαιώματος, επιτρέπεται η χρήση του περιεχομένου επί πληρωμή, ο δημιουργός απεκδύεται του περιουσιακού δικαιώματός του, όπως γίνεται και με τις άδειες τελικής χρήσης (EULA). Στις ανοικτές άδειες επιτρέπεται η περαιτέρω χρήση του περιεχομένου χωρίς οικονομικό τίμημα ή περιορισμούς ενώ ο δημιουργός διατηρεί το τεκμήριο κυριότητας, όπως γίνεται στις άδειες Creative Commons [10,13]. Μια άλλη διάκριση αφορά τις αποκλειστικές άδειες όπου η αδειοδότηση αφορά τον αποκλειστικό αδειολήπτη και τις ΜΗ αποκλειστικές όπου η αδειοδότηση μπορεί να αφορά πολλαπλούς αδειοδόχους. Η ανοιχτή αδειοδότηση οδηγεί σε άμεση διάθεση του περιεχομένου, αποσαφηνίζει τον δικαιούχο των δικαιωμάτων και ότι το έργο είναι προστατευμένο από λογοκλοπή, αναγράφει την ημερομηνία δημιουργίας και δημοσίευσης και περιγράφει τον τρόπο αναφοράς μέσω των Creative Commons (CC).

Τα CC, είναι μη κερδοσκοπικός οργανισμός για την προώθηση της ταυτοποίησης και επαναχρησιμοποίησης περιεχομένου με στόχο την καινοτομία. Οι άδειες CC αποτελούν το μέσο αναγνώρισης και προώθησης της πνευματικής ιδιοκτησίας για τους δημιουργούς με τη μορφή σύνθετης ετικέτας η οποία συνοδεύει το έργο και περιλαμβάνει λογότυπο και σύνδεσμο που οδηγεί στα μεταδεδομένα που συνοδεύουν το έργο.

Οι άδειες Creative Commons αποτελούν παράδειγμα ανοιχτής αδειοδότησης και είναι σχεδιασμένες για ανάγνωση από άνθρωπο ή μηχανή περιλαμβάνοντας στοιχεία αναγνώσιμα από

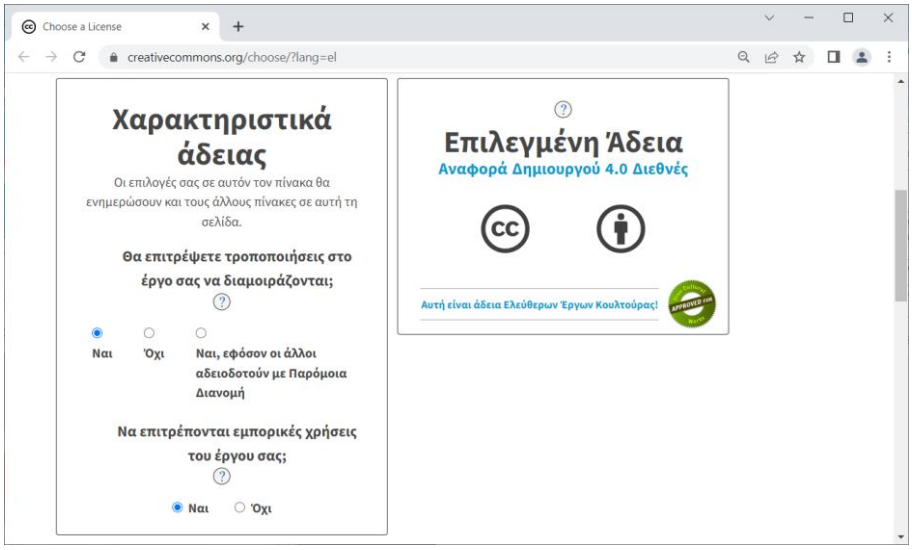
το ευρύ κοινό, τους νομικούς και τις μηχανές. Η ελληνική τεκμηρίωση της χρήσης των CC, υπάρχει στην ιστοσελίδα *[www.creativecommons.gr]*

Τα CC, περιλαμβάνουν διαφορετικούς τύπους αδειών που περιγράφονται από τα σχετικά ακρωνύμια όπως δείχνει ο Πίνακας 1.

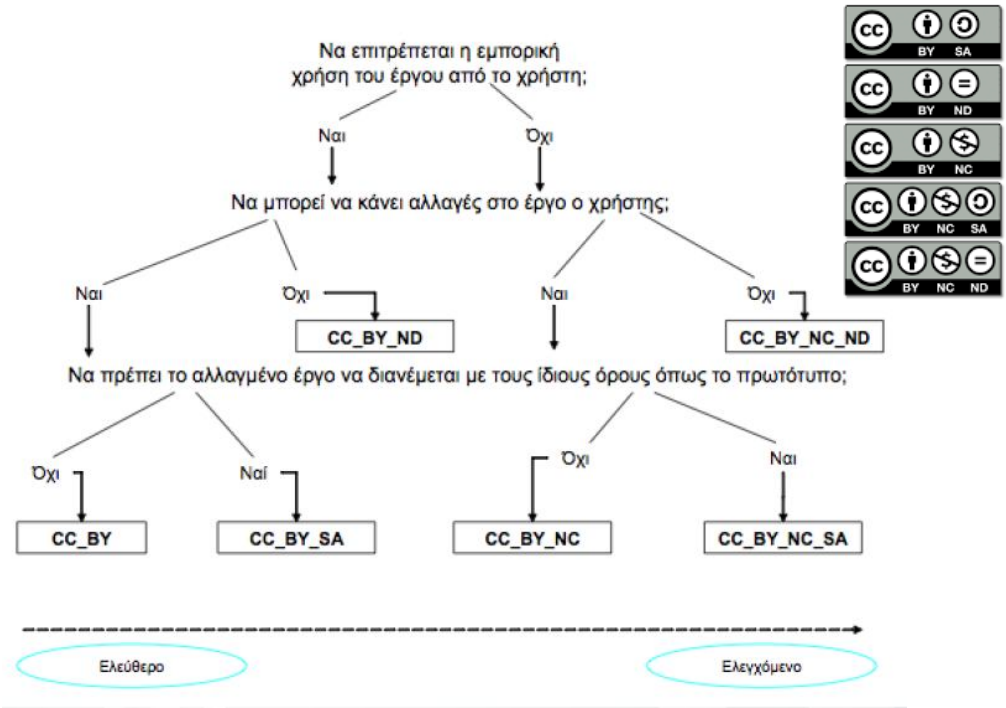
ΑΑ	***Τύπος άδειας***	***Όνομα***	***Περιγραφή***
1	*BY*	*Attribution*	*Αναφορά στον αρχικό δημιουργό σε περίπτωση αναπαραγωγής ή τροποποίησης σχετικού έργου*
2	*SA*	*Share Alike*	*Διανομή του παράγωγου έργου με τους όρους της αρχικής άδειας*
3	*NC*	*Non-Commercial*	*Απαγόρευση εμπορικής χρήσης του έργου*
4	*ND*	*No Derivative Works*	*Απαγόρευση δημιουργίας οποιουδήποτε είδους παραγώγων έργων*

Πίνακας 1: Τύποι αδειοδότησης Creative Commons

Η ιστοσελίδα *[www.creativecommons.gr]* παρέχει ένα σύντομο ερωτηματολόγιο που καθοδηγεί στη δημιουργία της κατάλληλης άδειας CC μέσα από έξι διαφορετικούς συνδυασμούς απαντήσεων [10,13].



Εικόνα 2: Αυτοματοποιημένη δημιουργία συνδυαστικής άδειας Creative Commons



Εικόνα 3: Δέντρο επιλογών στη δημιουργία σύνθετης αδειοδότησης Creative Common

Οι βάσεις δεδομένων, όπως αναφέρει η πηγή [11], είναι μέσο ανάπτυξης της οικονομίας αφού περιλαμβάνουν πληροφορίες απαιτούμενες στις αποφάσεις και στις αναλύσεις. Οι βάσεις δεδομένων έχουν μεταβληθεί σε πληροφορικά αγαθά μεγάλου μεγέθους και πολυπλοκότητας (BigData), μεγάλης αξίας και απολαμβάνουν νομικής προστασίας. Η οδηγία 96/9/ΕΟΚ και ο νόμος 2121/93, καθορίζουν τη βάση δεδομένων ως «συλλογή έργων, δεδομένων ή ανεξάρτητων στοιχείων, διευθετημένων κατά συστηματικό ή μεθοδικό τρόπο και ατομικώς προσιτών με ηλεκτρονικά ή μέσα ή κατ' άλλον τρόπο». Η δημιουργία και οργάνωση του υλικού μιας βάσης δεδομένων αποτελεί πρωτότυπο έργο και κατά συνέπεια πνευματικό δημιούργημα.

Ο δημιουργός μιας πρωτότυπης βάσης δεδομένων αποκτά αυτοδίκαια το δικαίωμα να ελέγχει την αναπαραγωγή και την επεξεργασία της βάσης. Οι ΜΗ πρωτότυπες ηλεκτρονικές βάσεις περιλαμβάνουν δομημένα πληροφορικά σύνολα (datasets) που εξυπηρετούν την έρευνα και αποτελούν επίσης αντικείμενο νομικής προστασίας.

Οι κατασκευαστές των ΜΗ πρωτότυπων συνόλων δεδομένων έχουν μια σειρά από αποκλειστικά δικαιώματα, όπως το να επιτρέπουν ή να αρνούνται την εξαγωγή δεδομένων, την αναπαραγωγή, την επαναχρησιμοποίηση του συνόλου ή ουσιώδους μέρους, να εκχωρούν δικαιώματα σε άλλους νόμιμους χρήστες οι οποίοι με τη σειρά τους μπορούν να εξάγουν και να χρησιμοποιούν επουσιώδη μέρη της βάσης δεδομένων για οποιονδήποτε σκοπό που δεν έρχεται σε σύγκρουση με τα οικονομικά συμφέροντα του κατασκευαστή.

Η νομική προστασία βάσεων δεδομένων διαφοροποιείται στις διαφορετικές χώρες. Το άρθρο [14] αναφέρει ότι οι βάσεις δεδομένων ως αποθήκες στατικών δεδομένων, αποτελούν τη βάση για το ηλεκτρονικό εμπόριο, το επιχειρείν, το μάρκετινγκ. Χωρίς πρόσβαση στην πληροφορία, δεν μπορεί να υπάρξει συμμετοχή στην εποχή της πληροφορίας. Παρόλο που οι βάσεις δεδομένων προστατεύονται από την παγκόσμια νομοθεσία περί πνευματικών δικαιωμάτων, ο ανταγωνισμός και το κυβερνο-έγκλημα συμμετέχουν σε πράξεις μη

εξουσιοδοτημένης πρόσβασης σε βάσεις δεδομένων και αντιγραφής τους με στόχο το παράνομο οικονομικό όφελος ή την απόκτηση ανταγωνιστικών πλεονεκτημάτων.

Στις ΗΠΑ και στην Ευρώπη, το οριζόμενο επίπεδο της πρωτοτυπίας έχει μεταβληθεί διαχρονικά. Στην Ινδία η προσέγγιση βασίζεται στη θεωρία της ανταμοιβής της επένδυσης σε εργασία, χρόνο και οικονομικούς πόρους. Έχουν προταθεί κατευθυντήριες προτάσεις για την ανάπτυξη νέας νομοθεσίας που θα συνυπολογίσει το ιδιωτικό και δημόσιο συμφέρον καθώς και το βασικό στόχο της πνευματικής ιδιοκτησίας που προωθεί τη δημιουργικότητα και την καινοτομία. Το άρθρο [14] αναφέρει μια σειρά από προτάσεις για την προστασία των βάσεων δεδομένων που περιλαμβάνουν την ανταπόδοση των επενδύσεων, τη διάκριση ιδιοκτησίας πρωτότυπων/ΜΗ πρωτότυπων, εμπορικών/ανοιχτών, ιδιωτικών/κυβερνητικών βάσεων δεδομένων, ένα ενιαίο σύστημα καταχώρησης δικαιωμάτων βάσεων δεδομένων, την επίβλεψη της εμπορικής εκμετάλλευσης από κυβερνητικές αρχές, την υποχρεωτικότητα στις αδειοδοτήσεις, την οριοθέτηση μικρότερης χρονικής περιόδου που καλύπτει η προστασία, την ενθάρρυνση της κοινωνίας των πολιτών για τη συνεισφορά τους στη δημιουργία της πληροφορίας.

Η εργασία [15] αναλύει την οπτική της πνευματικής κοινοκτημοσύνης στο σύγχρονο διαδίκτυο. Μια βάση ή συλλογή δεδομένων ορίζεται ως μια μεθοδικά οργανωμένη συλλογή στοιχείων ή έργων που αξιοποιεί ηλεκτρονικά μέσα για να διευκολύνει την αναζήτηση, την επεξεργασία την εξαγωγή δεδομένων και την επαναχρησιμοποίηση. Η εξαγωγή προσδιορίζεται νομικά στη μόνιμη ή προσωρινή μεταφορά του συνόλου ή ουσιώδους μέρους του περιεχομένου σε άλλο υλικό με οποιοδήποτε μέσο ή με οποιαδήποτε μορφή. Η επαναχρησιμοποίηση αντίστοιχα, περιγράφει την πάσης μορφής διάθεση στο κοινό του συνόλου ή ουσιώδους μέρους του περιεχομένου της βάσης δεδομένων με διανομή αντιγράφων, εκμίσθωση, μετάδοση. Η νομική προστασία του δημιουργού μιας βάσης δεδομένων σύμφωνα με το Αρ.3 της Ευρωπαϊκής Οδηγίας 96/9 δεν περιλαμβάνει και το περιεχόμενο ή τα επιμέρους έργα της συλλογής οπότε δεν υπάρχει σύγκρουση των δικαιωμάτων με τα πνευματικά δικαιώματα τρίτων.

Νόμιμος χρήστης είναι αυτός που έχει αποκτήσει το δικαίωμα χρήσης σε μια βάση δεδομένων και δύναται να εξάγει ή να επαναχρησιμοποιεί μόνο ασήμαντα μέρη του περιεχομένου, χωρίς να βλάπτει την εκμετάλλευση της βάσης από τον κατασκευαστή ή τα έννομα συμφέροντα του δημιουργού της. Το άρθρο [15] αναφέρει ότι το διαδίκτυο εξελίσσεται σε μια κοινωνία κοινής χρήσης πληροφοριών που περιλαμβάνει πνευματικά δημιουργήματα ενώ ταυτόχρονα συνυπάρχουν δεδομένα που απαιτούν εξουσιοδοτημένη πρόσβαση και δεδομένα που μπορούν να αγοραστούν. Το διαδίκτυο στηρίζει την ελεύθερη διακίνηση της πληροφορίας αλλά παράλληλα προσφέρει τα τεχνολογικά μέσα για την προστασία της πνευματικής ιδιοκτησίας. Ο όρος της πνευματικής ιδιοκτησίας περιγράφεται στο [15] ως «*κατοχή ενός άυλου αγαθού σε ένα οικονομικό σύστημα που βασίζεται στην ιδιοκτησία και στον έλεγχο των μέσων παραγωγής και ανταμείβει με βάση την παραγωγή*».

Οι θεωρίες που υποστηρίζουν ένα ανοιχτό σύστημα επιθυμούν μικρά κόστη δημιουργίας και διάδοσης που οδηγούν σε μεγαλύτερα κέρδη. Για παράδειγμα ένα ψηφιακό βιβλίο δεν απαιτεί χαρτί, εκτύπωση, μεταφορά, διάθεση και αποθήκευση σε βιβλιοπωλεία. Στον ψηφιακό κόσμο η δωρεάν διάθεση δεν δημιουργεί απώλειες αλλά δεν δημιουργεί ούτε προκαθορισμένο κέρδος. Στις περισσότερες χώρες του κόσμου τα πνευματικά έργα στο διαδίκτυο προστατεύονται με νομικά μέσα. Βέβαια η προστασία αυτή στην πράξη δεν επαρκεί αφού υπάρχουν πολλές περιπτώσεις

κατάχρησης. Αν κάποιος κατεβάσει ένα «βίντεοκλίπ», ηχογραφήσει ψηφιακά τον ήχο, διορθώσει την ποιότητά του με δωρεάν λογισμικό, θα αποκτήσει δωρεάν το έργο που επιθυμεί χωρίς να πληρώσει. Η πειρατεία δεν αφορά μόνο στο μοντέλο δωρεάν αναπαραγωγής και διάθεσης αλλά και στη ΜΗ εξουσιοδοτημένη πρόσβαση σε κλειστά κανάλια πρόσβασης.

Η πηγή [16] διακρίνει τα ανοιχτά δεδομένα, τα πνευματικά δικαιώματα και τα δεδομένα προσωπικού χαρακτήρα εξηγώντας τη διαδικασία της διάθεσης δεδομένων. Εκτενή αναφορά στα νομικά μέτρα προστασίας των ψηφιακών πνευματικών δικαιωμάτων γίνεται και στο [17].

Σε σχέση με τα προσωπικά δεδομένα, η εργασία [18] προσεγγίζει την προστασία των προσωπικών δεδομένων σε σχέση με την πνευματική ιδιοκτησία. Αναφέρεται στο θεσμικό πλαίσιο του ΓΚΠΔ και στο νομοθετικό πλαίσιο της Προστασίας της Πνευματικής Ιδιοκτησίας με τις ιδιομορφίες των πνευματικών δημιουργημάτων στον ανοιχτό χώρο του διαδικτύου. Αυτό καθιστά αναγκαία τη θέσπιση πρόσθετων κανόνων τονίζοντας την ιδιαιτερότητα των πνευματικών δικαιωμάτων.

Η εργασία [19], αναφέρεται στα συγγενικά δικαιώματα της πνευματικής ιδιοκτησίας που περιλαμβάνουν τα φυσικά η νομικά πρόσωπα εκτός του δημιουργού που συνεισφέρουν μέσα από την παρουσίαση των έργων ή την οικονομική επένδυση στην πνευματική δημιουργία, διαδραματίζοντας καθοριστικό ρόλο στη μεσολάβηση, στην παραγωγή, στη διάδοση και εκμετάλλευση των έργων. Τα συγγενικά δικαιώματα περιλαμβάνουν κάποιες περιορισμένες εξουσίες, συναφείς με αυτές των δημιουργών. Η εργασία αυτή επικεντρώνεται στο περιεχόμενο που δημιουργείται από εθελοντική εργασία (User Generated Content) και διαμοιράζεται μέσω κάποιας διαδικτυακής πλατφόρμας. Η άνοδος του UGC επηρεάζει την προστασία της προσωπικής ιδιοκτησίας και του Copyright. Οι πλατφόρμες UGC συνεργάζονται με ερασιτέχνες καλλιτέχνες, προβάλλοντας τα έργα τους. Πολλοί όμως χρήστες δημοσιεύουν υλικό που ανήκει σε άλλους δημιουργούς, αγνοώντας ή καταπατώντας τη νομοθεσία περί προστασίας των έργων αυτών. Οι κάτοχοι των πνευματικών δικαιωμάτων στη συνέχεια οδηγούνται σε νομικές ενέργειες και απαιτήσεις αποζημίωσης.

Η εργασία [51] ασχολείται με τα νομικά μέσα προστασίας βάσεων δεδομένων, εξετάζει τη νομοθεσία διαχρονικά με βάση τις Διεθνείς συνθήκες, το Ευρωπαϊκό δίκαιο (96/9/ΕΟΚ), και το εγχώριο δίκαιο ενώ αξιολογεί το αντίκτυπο των νομοθετικών ρυθμίσεων στην οικονομική ανάπτυξη. Αναλύει λεπτομέρειες της νομοθεσίας όπως ότι η προστασία επεκτείνεται στα στοιχεία που επιτρέπουν την αναζήτηση (θησαυρό, λεξικό, ευρετήριο), ότι η οδηγία παράλληλα με την ηθική αναγνώριση και την προστασία του δημιουργού, αναγνωρίζει το δικαίωμα της εμπορικής εκμετάλλευσης αποκλειστικά και μόνο στον τον κατασκευαστή που έχει επενδύσει στη δημιουργία της, ότι η έννοια του ουσιώδους δεν δίνεται στην Οδηγία, ούτε στο Ν. 2121/1993 αλλά δόθηκε από το Ευρωπαϊκό Δικαστήριο βάσει ποσοτικών μεγεθών όπως το ποσοστό των δεδομένων που έχει εξαχθεί από μια βάση ή έχει επαναχρησιμοποιηθεί σε σχέση με το συνολικό της όγκο. Η Ευρωπαϊκή Επιτροπή πραγματοποίησε μια on-line έρευνα και αξιολόγηση της Οδηγίας 96/9 στην ευρωπαϊκή βιομηχανία βάσεων δεδομένων (σε 500 εταιρείες) ενώ επικεντρώθηκε στο αντίκτυπο της νομοθεσίας στο ρυθμό ανάπτυξης της ευρωπαϊκής βιομηχανίας βάσεων δεδομένων. Τα συμπεράσματα αναφέρουν ότι ο οικονομικός αντίκτυπος στην παραγωγή βάσεων δεδομένων δεν απεδείχθη. Τα νομοθετικά κείμενα περιλαμβάνουν ασαφείς ορισμούς που οδηγούν τα εθνικά δικαστήρια σε ερμηνευτικές συνδρομές του Ευρωπαϊκού Δικαστηρίου. Η

προστασία που παρέχεται, είναι αντίθετη με την φιλοσοφία της ελεύθερης πρόσβασης και διάδοσης της πληροφορίας που επικρατεί σήμερα. Η ιδιοκτησιακή προσέγγιση των βάσεων δεδομένων φαίνεται να συγκρούεται με τη διαδικτυακή επανάσταση αφού επικεντρώνεται στην εμπορική διάθεση των δεδομένων. Πολλοί προτείνουν ως λύση, τις αναγκαστικές άδειες χρήσης, ως αντιστάθμισμα στον ιδιοκτησιακό χαρακτήρα που ελλοχεύει τον κίνδυνο μονοπώλησης κοινόχρηστων πληροφοριών. Η ίδια εργασία [51] κάνει εκτενή αναφορά στα νομικά μέτρα προστασίας της πνευματικής ιδιοκτησίας, ορίζοντας τις σχετικές έννοιες (υποκείμενο δικαιώματος, αντικείμενο προστασίας, περιεχόμενο του δικαιώματος, κατασκευαστής) και την εξέλιξη της νομοθεσίας αυτής σε διεθνές, ευρωπαϊκό και ελληνικό εθνικό επίπεδο. Η εργασία περιλαμβάνει μεγάλη ελληνική βιβλιογραφία σε σχέση με το νομικό πλαίσιο που ισχύει σήμερα.

2.2 Σχετικά με τις τεχνικές υδατογράφησης δεδομένων

Η βιβλιογραφία σχετικά με τις τεχνικές υδατοσήμανσης σχεσιακών δεδομένων είναι μεγάλη και σε αυτή την ενότητα συνοψίζονται διαφορετικές υλοποιήσεις και έρευνες που κατηγοριοποιούν και συγκρίνουν τα διαφορετικά σχήματα.

Το άρθρο [20] εξηγεί την ανάγκη της προστασίας πνευματικών δικαιωμάτων βάσεων δεδομένων με σύγχρονα τεχνολογικά μέσα. Αναφέρει την αρχαία μεθοδολογία εισαγωγής σφαλμάτων σε πίνακες για τον εντοπισμό πειρατικών αντιγράφων ενώ εξηγεί ένα σύγχρονο αλγόριθμο υδατογράφησης που χρησιμοποιεί το μοντέλο ενός σύννεφου που αποτελείται από σωματίδια με τυχαία διασπορά. Ο αλγόριθμος προσθήκης και ανίχνευσης υλοποιείται και δοκιμάζεται μέσα από πειράματα. Η μέθοδος υδατογράφησης αξιοποιεί αριθμητικά γνωρίσματα βάσης δεδομένων, στα οποία γίνονται μικρές αλλαγές των τιμών ώστε να μην επηρεάζεται η χρηστικότητα των δεδομένων. Οι αλλαγές βασίζονται σε ένα μοτίβο από bits που έχει δημιουργηθεί εκ των προτέρων και μοιάζει με τη συγκέντρωση υδρατμών σε σύννεφο. Τα κριτήρια της αξιολόγησης της προσέγγισης, αφορούν τις γενικότερες απαιτήσεις ενός υδατογραφήματος που είναι η αδιορατότητα (να μην είναι ορατό), η διακριτικότητα (μη καταστροφή των στοιχείων), η ανθεκτικότητα (σε επιθέσεις), η δυνατότητα προσθήκης και ανίχνευσης, η χαμηλού κόστους υλοποίηση. Η έρευνα προτείνει την επέκταση σε μη αριθμητικά δεδομένα.

Η εργασία [21] αναλύει μια πολύ δημοφιλή τεχνική για την ισχυρή υδατογράφηση βάσεων δεδομένων με απώτερο σκοπό την αναγνώριση του δημιουργού και την αποτροπή της πειρατείας δεδομένων. Αναλύει τις απαιτήσεις μιας υδατοσήμανσης για σχεσιακά δεδομένα και παρουσιάζει τον αλγόριθμο δημιουργίας και ανίχνευσης υδατογραφήματος. Η τεχνική δημιουργίας μεταβάλει bits αριθμητικών δεδομένων προσθέτοντας σημάδια (marks) μέσα από ένα μοτίβο (watermark) που καθορίζεται αλγοριθμικά υπό τον έλεγχο ενός μυστικού κλειδιού και μιας σειράς παραμέτρων που αφορούν τον αριθμό των επηρεαζόμενων bits και τον αριθμό των επηρεαζόμενων πλειάδων. Η ανίχνευση γίνεται με τη χρήση του ίδιου κλειδιού στα υδατογραφημένα δεδομένα, χωρίς πρόσβαση στα αρχικά δεδομένα. Η ανάλυση δείχνει ότι η προτεινόμενη τεχνική είναι αποδοτική και ανθεκτική σε διάφορες μορφές μεταβολών αλλά και κακόβουλων επιθέσεων. Στο ίδιο άρθρο [21] αναφέρονται οι διάφοροι τύποι επιθέσεων: καλοήθεις ενημερώσεις, διαγραφές ή εισαγωγές πλειάδων, αλλαγές bits αριθμητικών τιμών σε 0 ή 1 με τυχαίο η καθοριστικό τρόπο,

στρογγυλοποίηση τιμών αριθμητικών χαρακτηριστικών, γραμμικοί μετασχηματισμοί αριθμητικών τιμών, εξαγωγή υποσυνόλου (πλειάδων ή ιδιοτήτων), ψεύτικοι ισχυρισμοί ιδιοκτησίας όπου ο επιτιθέμενος εισάγει δικό του υδατογράφημα και αξιώνει ιδιοκτησία. Ένα αποτελεσματικό σύστημα υδατογράφησης πρέπει να πληροί τις ιδιότητες της ανθεκτικότητα σε ενημερώσεις ή επιθέσεις, της ακρίβειας στην ανίχνευση, τη δυνατότητα αυξητικής προσθήκης, την τυφλή προσθήκη/ανίχνευση χωρίς πρότερη γνώση της αρχικής έκδοσης, τη χρήση δημόσιων αλγόριθμων (κατά Kerckhoff) και ιδιωτικών κλειδιών στην επίτευξη της ασφάλειας. Η συνεισφορά αυτής της εργασίας [21] είναι η οριοθέτηση των επιθυμητών ιδιοτήτων μιας υδατοσήμανσης, η τεκμηρίωση των πιθανών επιθέσεων σε ένα υδατογράφημα ΒΔ, η περιγραφή, υλοποίηση και αξιολόγηση μιας τεχνικής υδατογράφησης μέσα από πειράματα με σύνολα δεδομένων. Η έρευνα προτείνει την επέκταση σε μη αριθμητικά δεδομένα.

Το άρθρο [22] απασχολείται με την ψηφιακή υδατοσήμανση δεδομένων ως αποδοτικό τεχνικό μέτρο κατά της ψηφιακής πειρατείας. Η διαδικασία περιγράφει έναν νέο σχήμα υδατογράφησης για σχεσιακά δεδομένα που περιέχουν αριθμητικά στοιχεία, ένα ειδικό αλγόριθμο για υδατογραφήματα που λειτουργούν ως σφραγίδα στον προσδιορισμό του πνευματικού ιδιοκτήτη. Η εργασία παρέχει πειραματικά αποτελέσματα στη δημιουργία, απόδοση και ανθεκτικότητα σε πιθανές επιθέσεις. Τα σημάδια της υδατογράφησης εφαρμόζονται από έναν κωδικοποιητή ενώ η ανίχνευση γίνεται με τη χρήση ιδιωτικού κλειδιού. Το προτεινόμενο σχήμα μπορεί να χρησιμοποιηθεί για υδατογράφηση και ως δακτυλικό αποτύπωμα. Ο αλγόριθμος κωδικοποίησης μπορεί να εφαρμοστεί σε κάθε πλειάδα ανεξάρτητα, οπότε η προτεινόμενη μέθοδος έχει την ιδιότητα της σταδιακής δυνατότητα ενημέρωσης, δηλαδή η υδατογραφημένη βάση δεδομένων μπορεί να υποστηρίξει κανονικές τροποποιήσεις χρήστη (εισαγωγές, διαγραφές και ενημερώσεις) εφαρμόζοντας απλώς τον αλγόριθμο κωδικοποίησης σε επίπεδο πλειάδας και όχι στο σύνολο του πίνακα. Ο αλγόριθμος ταξινομεί τα bit κάθε πλειάδας σε μια μυστική σειρά και επιλέγει μερικά από τα bit δεδομένων του για να δρομολογήσει την πλειάδα σε ένα συγκεκριμένο bit υδατογραφήματος και ένα bit δεδομένων που θα επισημανθεί από την τιμή του εκχωρημένου bit υδατογραφήματος. Η έρευνα προτείνει να επεκταθεί και σε μη αριθμητικά δεδομένα, δηλαδή σε κατηγορικές και αλφαριθμητικές ιδιότητες.

Μια άλλη εργασία [23] προτείνει δύο σχήματα ενσωμάτωσης, μια υδατοσήμανση και ένα αποτύπωμα για την προστασία από παράνομη αναδιανομή αλλά και αναπαραγωγή. Το αποτύπωμα είναι ενσωματωμένο για κάθε παραλήπτη στον οποίο διανέμεται η βάση δεδομένων. Τα πειράματα έδειξαν ότι το προτεινόμενο σύστημα είναι αποτελεσματικό και ανθεκτικό σε διάφορες επιθέσεις. Ένα ολοκληρωμένο σύστημα δοκιμών αλγορίθμων υδατογράφησης ΒΔ προτείνεται στο άρθρο [24] με το όνομα «dbMark». Το εργαλείο συνοδεύεται από ένα σύνολο τριών υλοποιημένων σχημάτων υδατογράφησης που μπορούν να χρησιμοποιηθούν άμεσα από μελετητές για εκπαιδευτικούς σκοπούς. Το σετ περιέχει την εφαρμογή μιας μεθόδου ανά κατηγορία υδατοσήμανσης σχήματος : ανθεκτικό, εύθραυστο και δακτυλικό αποτύπωμα.

Όσον αφορά την ισχυρή υδατοσήμανση, το dbMark παρέχει την υλοποίηση του συστήματος υδατογράφησης των Agrawal και Kiernan [21], το οποίο είναι το πιο γνωστό ισχυρό σχήμα στη βιβλιογραφία. Όσον αφορά την εύθραυστη υδατοσήμανση, το πιο διάσημο σχήμα [58] έχει επίσης εφαρμοστεί και είναι έτοιμο για χρήση. Τέλος, όσον αφορά τη υδατοσήμανση δακτυλικών αποτυπωμάτων, το dbMark περιλαμβάνει μια υλοποίηση της παραμόρφωσης που βασίζεται στη μέθοδο [22] για αριθμητικές τιμές ως εναλλακτική λύση δακτυλικών

αποτυπωμάτων. Το dbMark διαθέτει διεπαφή χρήστη, είναι παραμετρικό και ευέλικτο στη συγκριτική αξιολόγηση μεθόδων υδατογράφησης για σχεσιακές βάσεις δεδομένων, μέσω της υλοποίησης επιθέσεων και της μέτρησης δεικτών αξιολόγησης. Ο χρήστης του εργαλείου μπορεί για παράδειγμα να επιλέξει αριθμητικά χαρακτηριστικά της βάσης MySQL στα οποία να επιτεθεί καθώς και να ενσωματώσει νέες υλοποιήσεις υδατογράφησης (σε Java), επιθέσεων και μετρικές αξιολόγησης.

Η εργασία [25] προτείνει μια μέθοδο, βασισμένη στην εικόνα ως υδατογράφημα, ενσωμάτωσης τη υδατοσήμανσης στη βάση δεδομένων σε δύο διαφορετικά χαρακτηριστικά πλειάδας, σε αριθμητικό πεδίο και στην ώρα πεδίου ημερομηνίας. Υπάρχουν πολλές προσεγγίσεις υδατοσήμανσης ΒΔ που περιλαμβάνουν την ενσωμάτωση ήχου, εικόνας και ΜΗ αριθμητικών πεδίων με την προσθήκη κενών χαρακτήρων. Η προτεινόμενη προσέγγιση ενσωματώνει το ίδιο υδατογράφημα σε δύο χαρακτηριστικά ταυτόχρονα, δυσκολεύοντας την επίθεση. Η διαδικασία της ενσωμάτωσης και της εξαγωγής του εκτελείται σε δύο φάσεις, στην πρώτη φάση εισάγεται το υδατογράφημα σε αριθμητικό πεδίο της βάσης δεδομένων και στη δεύτερη φάση εισάγεται ένα υδατογράφημα στο πεδίο seconds της βάσης δεδομένων. Κατά την εξαγωγή, ακολουθείται η αντίστροφη σειρά των φάσεων της ενσωμάτωσης. Η μέθοδος αλλάζει πολλά χαρακτηριστικά μιας πλειάδας που είναι απαραίτητο για μια ισχυρή προστασία πνευματικών.

Η δημοσίευση [26] αφορά μια πρόσφατη εκτεταμένη έρευνα με ταξινόμηση των μεθόδων υδατογράφησης ενώ περιλαμβάνει χρήσιμα διαγράμματα και 129 βιβλιογραφικές πηγές.

Το άρθρο [27], προτείνει ένα αναστρέψιμο αλγόριθμο υδατοσήμανσης βάσης δεδομένων που χρησιμοποιεί τον αλγόριθμο βελτιστοποίησης Firefly, εμπνευσμένο από τον κόσμο της βιολογίας και συγκεκριμένα των πυγολαμπίδων. Το βασικό πλεονέκτημα του αλγόριθμου είναι η εύκολη εφαρμογή του και η αποδοτικοτητά του. Η μέθοδος περιγράφει την εισαγωγή και την εξαγωγή του υδατογραφήματος. Η εξαγωγή, αφαιρεί το υδατογράφημα και αναστρέφει τη βάση στην αρχική της μορφή. Η εργασία περιλαμβάνει πειράματα ανθεκτικότητας σε επιθέσεις και μετρήσεις αποτελεσματικότητας.

Μια μελέτη τεχνικών για σχεσιακές βάσεις δεδομένων [28] αναλύει τη χρησιμότητα της ψηφιακής υδατοσήμανσης στις σχεσιακές βάσεις δεδομένων ως λύση για την προστασία των πνευματικών δικαιωμάτων, τον εντοπισμό παραβιάσεων, τον εντοπισμού πειρατών και τη διατήρηση ακεραιότητα των σχεσιακών δεδομένων. Εξηγεί τη διαδικασία της δημιουργίας και επαλήθευσης μιας υδατοσήμανσης, τα είδη επιθέσεων και τις τεχνικές άμυνας. Στη συνέχεια εξετάζει όλες τις γνωστές τεχνικές υδατογράφησης για βάσεις δεδομένων (έως το 2010) που έχουν προταθεί στη βιβλιογραφία για την αντιμετώπιση διαφορετικών στόχων και τις ταξινομεί σε συνοπτικό πίνακα με βάση το σκοπό, τον τρόπο δημιουργίας, τους τύπους δεδομένων όπου εφαρμόζονται, το επίπεδο λεπτομέρειας και την επαληθευσιμότητα του υδατογραφήματος που παράγεται. Καταλήγει στο συμπέρασμα οτι οι περισσότερες από τις τεχνικές που βασίζονται σε παραμόρφωση στοχεύουν κυρίως στην προστασία της ιδιοκτησίας, ενώ η υδατοσήμανση χωρίς παραμόρφωση είναι εύθραυστη και στοχεύει στη διατήρηση της ακεραιότητας της πληροφορίας.

Το άρθρο [29] αφορά την υδατοσήμανση δεδομένων XML που έχουν την ιδιαιτερότητα ότι η προσθήκη θορύβου μπορεί να καταστρέψει την ιεραρχική τους δομή, σε αντίθεση με τα πολυμέσα που έχουν υψηλή ανοχή στο θόρυβο. Η εργασία παρουσιάζει μεθόδους που ενσωματώνουν θόρυβο, διατηρώντας αναλλοίωτη τη δομή. Το σχήμα υδατογράφησης ακολουθεί

το σκεπτικό του αλγόριθμου υδατογράφησης δεδομένων των Agrawal & Kiernan [21]. Η δεύτερη προσέγγιση προσθέτει υδατοσήμανση σε συμπιεσμένη μορφή του αρχείου. Παρουσιάζονται πειράματα με επιθέσεις και ελέγχεται η αποτελεσματικότητα και η αποδοτικότητα της προσέγγισης.

Η δημοσίευση [30], υλοποιεί μέθοδο υδατοσήμανσης βάσεων δεδομένων σε Java, παρεμβαίνοντας σε αριθμητικές τιμές και τη δοκιμάζει σε μια εμπορική βάση δεδομένων πωλήσεων της Wal-Mart. Περιγράφεται η διαδικασία καθορισμού του επιπέδου παρέμβασης με βάση τη στατιστική ανάλυση των τιμών της βάσης, οι αλγόριθμοι δημιουργίας και ανίχνευσης, τα πειράματα επιθέσεων και αποτελεσματικότητας με διαγράμματα.

Το άρθρο[31] προτείνει ένα σχήμα υδατοσήμανσης που δεν εξαρτάται από τα πρωτεύοντα κλειδιά τα οποία μπορεί να αλλοιωθούν σε μια επίθεση. Το νέο σχήμα κατασκευάζει εικονικά πρωτεύοντα κλειδιά από τα πλέον σημαντικά bits κάποιων γνωρισμάτων όπου η επιλογή των γνωρισμάτων βασίζεται σε ένα μυστικό κλειδί. Τα εικονικά πρωτεύοντα κλειδιά χρησιμοποιούνται στη δημιουργία της υδατοσήμανσης. Τα πειράματα απέδειξαν τη μεγάλη ανθεκτικότητα των υδατοσημάτων σε διάφορες επιθέσεις. Επίσης το άρθρο [32] περιγράφει ένα αλγόριθμο υδατοσήμανσης που χωρίζει τα δεδομένα σε ομάδες με τη χρήση συναρτήσεων κατακερματισμού, προσθέτει μικρό θόρυβο στα αριθμητικά δεδομένα των επιλεγμένων ομάδων παρεμβαίνοντας σε bits δεδομένων με βάση ένα προϋπολογισμένο επίπεδο θορύβου.

Στην εργασία [34], παρουσιάζεται ένα σύστημα υδατογράφησης για αριθμητικές βάσεις δεδομένων, που διατηρούν τις στατιστικές ιδιότητες των τιμών τους (μέση τιμή και διακύμανση). Η προσέγγιση περιλαμβάνει δημιουργία και ανίχνευση που αποφασίζει εάν ένα υδατογράφημα είναι ενσωματωμένο ή όχι. Οι πιθανότητες ψευδώς θετικών και ψευδώς αρνητικών αποτελεσμάτων μπορούν να περιοριστούν με βάση μια παράμετρο. Το προτεινόμενο σύστημα φαίνεται ανθεκτικό στην προσθήκη θορύβου μέσα από πειράματα.

Μια πρόσφατη εργασία (2022) [35] προτείνει ένα ιεραρχικό σύστημα προστασίας της πνευματικής ιδιοκτησίας συνόλων δεδομένων που βασίζονται σε διπλή υδατοσήμανση χρησιμοποιώντας μηχανική μάθηση. Το σύστημα διαχωρίζει τους εσωτερικούς χρήστες από τους συνδρομητές και παρέχει υπηρεσίες ελέγχου υδατοσήμανσης με τη χρήση εικόνων. Το άρθρο περιγράφει πειράματα σε σύνολα δεδομένων για να επαληθεύσει την αποτελεσματικότητα τη μεθόδου. Το άρθρο [36] παρουσιάζει μια εκτεταμένη βιβλιογραφική έρευνα αναστρέψιμων μεθόδων υδατογράφησης. Η ταξινόμηση βασιζεται στο μέγεθος των παρεμβάσεων στα δεδομένα και στην ανθεκτικότητα απέναντι σε κακόβουλες επιθέσεις. Οι αναστρέψιμες υδατογραφήσεις πέρα από την κωδικοποίηση (προσθήκη) και την αποκωδικοποίηση (ανίχνευση) των υδατογραφημάτων, περιλαμβάνουν και μια τρίτη φάση που επαναφέρει τα δεδομένα στην αρχική τους μορφή αφαιρώντας το υδατογράφημα. Οι απαιτήσεις των αναστρέψιμων μεθόδων περιλαμβάνουν τη δυνατότητα επαναφοράς, την ελάχιστη παρέμβαση, την τυφλή υδατογράφηση, την ανθεκτικότητα, την ασφάλεια. Η έρευνα συγκεντρώνει όλα τα γνωστά ισχυρά και εύθραυστα σχήματα υδατογράφησης σε πίνακες και καταγράφει επιπρόσθετα και τα αναστρέψιμα. Επίσης περιλαμβάνει μια συλλογή από 81 πηγές, με όλα τα γνωστά στο χώρο σχήματα υδατογράφησης αριθμητικών και ΜΗ, δεδομένων.

Μια κατηγορία άρθρων αφορά την υδατογράφηση ανωνυμοποιημένων δεδομένων [38,39]. Το άρθρο [38] εφαρμόζει τεχνικές υδατοσήμανσης σε ανωνυμοποιημένα δεδομένα, με στόχο να

αποτρέψει τη μη εξουσιοδοτημένη αναδιανομή ανώνυμων δεδομένων. Η εργασία προτείνει μια μέθοδο υδατογράφησης που ενσωματώνει συμβολοσειρές από bits σε ανώνυμα δεδομένα κατά τη δημοσίευση με τη χρήση συμμετρικής κρυπτογράφησης AES. Για να μειωθεί η απώλεια πληροφοριών, επιλέγει τιμές αντικατάστασης των οποίων η σημασία είναι παρόμοια με την αρχική. Τα πειράματα έδειξαν ότι η διαδικασία οδηγεί σε μικρή απώλεια της πληροφορίας και είναι ανθεκτική σε επιθέσεις. Η εργασία [39] προτείνει μια μέθοδο υδατογράφησης για ανώνυμα δεδομένα με στόχο τον εντοπισμό του πειρατή που αναδημοσιεύει ανώνυμα δεδομένα χωρίς εξουσιοδότηση. Η μέθοδος υδατογράφησης αξιοποιεί το συμμετρικό αλγόριθμο κρυπτογραφίας AES για την προστασία του υδατογραφήματος από επιθέσεις παραμόρφωσης. Τα αποτελέσματα των πειραμάτων έδειξαν ότι η προτεινόμενη υδατοσήμανση μπορεί να εξάγει περισσότερο από το 95% των υδατογραφημένων ανώνυμων δεδομένων. Επίσης, η υδατοσήμανση μπορεί να ανιχνευτεί και από ένα υποσύνολο (30%) των ανώνυμων δεδομένων.

Πολλές τεχνικές υδατογράφησης έχουν προταθεί στη βιβλιογραφία που συνήθως περιορίζονται σε αριθμητικά χαρακτηριστικά. Αυτές οι τεχνικές δεν είναι αποτελεσματικές στη διαχείριση μικρών αριθμών. Η δημοσίευση [40] προτείνει μια υδατοσήμανση όπου το τρίτο πιο σημαντικό bit αριθμητικών αλλά και μη αριθμητικών δεδομένων αντικαθίσταται με το bit του υδατογραφήματος χωρίς σημαντική απώλεια πληροφορίας. Η πρόταση έχει υλοποιηθεί και ελεγθεί σε πειράματα ανθεκτικότητας ενώ κρίνεται κατάλληλη για εφαρμογή σε εμπορικές τράπεζες όπου εμφανίζονται συχνά περιπτώσεις απάτης. Επίσης το άρθρο [44], ασχολείται με την υδατογράφηση αλφαριθμητικών αραβικών χαρακτηριστικών σε βάσεις δεδομένων. Πολλοί αραβικοί χαρακτήρες είναι επεκτάσιμοι και επιτρέπουν την προσθήκη bits στις προεκτάσεις τους χωρίς να αλλοιώνεται η αναγνωσιμότητα των χαρακτήρων. Το (ένα) μπορεί να γραφτεί (ﺳ ٘ واح) ή σε εκτεταμένη μορφή (ﺩ واح) ενώ η δεύτερη έκδοση μπορεί να φέρει πρόσθετες δυαδικές πληροφορίες. Στον αλγόριθμο, γίνεται χρήση μιας εικόνας η οποία διαιρείται σε τμήματα (δυαδικές συμβολοσειρές) που κωδικοποιούνται σε μη αριθμητικά χαρακτηριστικά πολλαπλών πλειάδων της βάσης δεδομένων. Ένα πλεονέκτημα της προσέγγισης είναι η διαθέσιμη χωρητικότητα στην απόκρυψη μεγάλων υδατογραφημάτων. Η ανθεκτικότητα του αλγόριθμου δοκιμάσθηκε με επιθέσεις διαγραφής, προσθήκης και αλλαγής υποσυνόλου.

Η έρευνα [41], αποτελεί μια μεγάλη συγκέντρωση στοιχείων από 113 βιβλιογραφικές πηγές για την ταξινόμηση όλων των γνωστών μεθόδων υδατοσήμανσης σε πολυμέσα, εικόνες, ήχο, κείμενο και βάσεις δεδομένων. Εξηγεί όλες τις μορφές επιθέσεων, τους δείκτες αποδοτικότητας και τις γνωστές μετρικές απόδοσης. Η συλλογή και αξιολόγηση των τεχνικών έχει στόχο να βοηθήσει τους μελλοντικούς ερευνητές στο να επιλέξουν την καταλληλότερη μέθοδο με βάση τον σκοπό της υδατοσήμανσης.

Σε σχέση με το Copyrighting as a Service (CaaS), το άρθρο [42] περιγράφει μια ηλεκτρονική υπηρεσία ψηφιακής υδατογράφησης που υλοποιεί ένα μοντέλο με τρία υδατοσημάτα, του δημοσίου, του ιδιοκτήτη και του αποδέκτη (συνδρομητή) με στόχο την ιχνηλάτηση των ψηφιακών έργων. Η εργασία δεν αφορά δεδομένα αλλά αρχεία πολυμέσων, όμως έχει ενδιαφέρον γιατί το μοντέλο που προτείνει παρέχει την υδατογράφηση ως υπηρεσία (WaaS). Αντίστοιχα η εργασία [43] προτείνει την ανώνυμη υδατογράφηση ως υπηρεσία (WaaS) χρησιμοποιώντας μηδενικά σχήματα υδατοσήμανσης. Η τεχνική δεν προσθέτει επιπλέον πληροφορία στη βάση δεδομένων. Η εξαγωγή γίνεται με υψηλή πιθανότητα επιτυχίας που αποδεικνύει την ανθεκτικότητα σε επιθέσεις. Το σχήμα της υδατογράφησης βασίζεται στην

προσέγγιση της μηδενικής υδατοσήμανσης όπου ο αλγόριθμος ενσωμάτωσης δεν προσθέτει μυστική πληροφορία στα δεδομένα. Το υδατογράφημα δημιουργείται από χαρακτηριστικά της βάσης και καταχωρείται σε μια κεντρική Τρίτη Αρχή Πιστοποίησης. Αυτή η τεχνική διατηρεί την ποιότητα της βάσης δεδομένων και μπορεί να εφαρμοστεί για κάθε τύπο βάσης με αριθμητικές ή αλφαριθμητικές τιμές.

Η εργασία [46] ασχολείται με την υδατοσήμανση δεδομένων που διατηρούν τη σημασιολογική αξία τους ανέπαφη, επιστρέφοντας τα ίδια αποτελέσματα σε ερωτήματα του χρήστη. Η διαδικασία χρησιμοποιεί αλφαριθμητικά δεδομένα. Η επιλογή των πλειάδων και των χαρακτηριστικών γίνεται χρησιμοποιώντας ένα μυστικό κλειδί ενώ το ίδιο κλειδί χρησιμοποιείται στην ανίχνευση χωρίς να χρειάζεται η αρχική μορφή (τυφλό σχήμα). Η μέθοδος αποδείχθηκε ανθεκτική απέναντι σε γνωστές επιθέσεις και αποτελεσματική για χρήση σε πραγματικά δεδομένα.

Το άρθρο [47] περιγράφει τη νομοθεσία (σε ΗΠΑ, Ευρώπη, Ινδία), τα τεχνικά χαρακτηριστικά της υδατογράφησης, τις απαιτήσεις ενός σχήματος υδατογράφησης, τις τεχνικές υδατογράφησης τις μορφές επιθέσεων και περιλαμβάνει συγκεντρωτικούς πίνακες με όλες τις γνωστές υλοποιήσεις και τις αδυναμίες τους. Επιπρόσθετα η μελέτη περιλαμβάνει προσεγγίσεις σε μη αριθμητικούς τύπους δεδομένων (κείμενο, ώρα, κατηγορία). Ολοκληρώνει με τις μελλοντικές κατευθύνσεις της έρευνας που αφορούν, την υδατοσήμανση web databases, no-SQL, object oriented με ενσωματωμένα αντικείμενα πολυμέσεων. Επίσης αναφέρει τη χρήση βιομετρικών στοιχείων (φωτογραφιών, ήχου) στην αυθεντικοποίηση του δημιουργού.

Η συλλογή άρθρων [48] αγγίζει διάφορα θέματα κυβερνοασφάλειας. Το πρώτο άρθρο παρουσιάζει ένα νέο ισχυρό σχήμα υδατοσήμανσης δεδομένων που προσαρμόζεται με βάση τους τύπους δεδομένων των χαρακτηριστικών (αριθμητικά ή κείμενο) και υλοποιεί διάφορες τεχνικές της βιβλιογραφίας ανάλογα με την περίπτωση. Στα αριθμητικά δεδομένα, χρησιμοποιείται ο αλγόριθμος τροποποίησης LSB και η αναζήτηση μοτίβων. Στα δεδομένα κειμένου, ο αλγόριθμος ενσωμάτωσης χώρου, τροποποίησης συμβόλων και λεξικά. Ο αλγόριθμος προσαρμόζεται αυτόματα με βάση τους τύπους δεδομένων, τον όγκο δεδομένων, την ευαισθησία δεδομένων ενώ υλοποιεί τεχνικές ανάλυσης, αυτόματης ρύθμισης παραμέτρων και οπτικοποίησης. Η απόδοση και η αποτελεσματικότητα έχουν επαληθευτεί πειραματικά απέναντι στις πιο συνηθισμένες επιθέσεις.

Η εργασία [49] προτείνει μια αποτελεσματική τεχνική για την υδατοσήμανση βάσης δεδομένων όπου μια κατάλληλη πλειάδα επιλέγεται για σήμανση και τα επιλεγμένα bits μιας εικόνας αντικαθιστούν μερικά bit των επιλεγμένων χαρακτηριστικών της πλειάδας. Η τεχνική ελαχιστοποιεί τη διακύμανση αναστρέφοντας μερικά κομμάτια του υδατογραφημένου χαρακτηριστικού. Τα πειραματικά αποτελέσματα απέδειξαν την ανθεκτικότητα της τεχνικής, έναντι σε επιθέσεις διαγραφής, μεταβολής, αλλαγής σειράς και εισαγωγής πλειάδων.

Η διδακτορική διατριβή [50] αφορά τη μελέτη της υδατοσήμανσης δεδομένων υγείας με στόχο τη φροντίδα ασθενών, τη διαχείριση νοσοκομείου ή τη δημόσια υγεία. Οι βάσεις δεδομένων μπορούν να μεταφερθούν και να μοιραστούν μεταξύ διαφορετικών ιδρυμάτων υγείας ή να τροφοδοτήσουν αποθετήρια. Επισημαίνεται οτι η υδατοσήμανση ως μορφή κρυπτογράφησης με ελεγχόμενη ενσωμάτωση θορύβου, επιτρέπει την πρόσβαση στα δεδομένα διατηρώντας τα παράλληλα προστατευμένα. Η εργασία εξετάζει την ιχνηλασιμότητα βάσεων δεδομένων που συγχωνεύονται σε κοινόχρηστες αποθήκες δεδομένων (data warehouses) και προτείνει μια

στρατηγική ανίχνευσης μιας βάσης δεδομένων που αποτελεί το 7% ενός μίγματος, με σχεδόν απόλυτη επιτυχία. Η ασφάλεια του συστήματος ακολουθεί τις αρχές της κρυπτογραφίας που υλοποιούν ανοιχτούς αλγορίθμους και μυστικά κλειδιά. Η στρέβλωση που εισάγεται στη βάση δεδομένων αξιολογείται στατιστικά και σημασιολογικά όπου ο όρος «στατιστική» αναφέρεται στη διασπορά των αλλαγών στο σύνολο των πλειάδων. Στις βάσεις δεδομένων υγείας υπάρχουν ισχυροί σημασιολογικοί δεσμοί ανάμεσα στις τιμές των χαρακτηριστικών, οπότε η εισαγωγή θορύβου μπορεί να χαλάσει αυτούς τους δεσμούς.

Η έρευνα-παρουσίαση [52] εξερευνά την απόκρυψη πληροφορίας ως εργαλείο στην προστασία δικαιωμάτων βάσεων δεδομένων σε αριθμητικά και κατηγοριοποιημένα δεδομένα. Διαφοροποιεί την υδατοσήμανση-watermark από την απόκρυψη-steganography με βάση το σκοπό και όχι την τεχνική. Η απόκρυψη στοχεύει στην ανταλλαγή ανθεκτικών και κρυφών μηνυμάτων ενώ η υδατογράφηση ως απόδειξη των δικαιωμάτων επί του πνευματικού έργου με έμφαση στην ανίχνευση και όχι στην εξαγωγή. Εάν η ανάκτηση των δεδομένων έχει μεγαλύτερη αξία από την ανίχνευση, αυτό υποδηλώνει τους στόχους της στεγανογραφίας. Επίσης η πηγή, εξηγεί τους δημοφιλέστερους αλγόριθμους single bit και multi bit και τις μορφές επιθέσεων.

Η μεταπτυχιακή διατριβή [53], αναπτύσσει μια τεχνική υδατογραφήματος για σχεσιακή βάσεις δεδομένων με στόχο την αναγνώριση του ιδιοκτήτη και τον εντοπισμό πειρατών μέσα από ένα μοντέλο όπου συνδρομητές αγοράζουν δεδομένα. Πραγματοποιεί πειράματα ανθεκτικότητας ενώ παράλληλα ασχολείται με την αντιστροφή της υδατοσήμανσης για την ανάκτηση των αρχικών δεδομένων. Η εργασία εισάγει την ιδέα της ιεραρχικής πρόσβασης στα δεδομένα και της αυξητικής υδατοσήμανσης όταν τροποποιείται ή ενημερώνεται ένα μικρό τμήμα της βάσης δεδομένων.

Η εργασία [58] αφορά την ανίχνευση μη εξουσιοδοτημένων τροποποιήσεων σε βάσεις δεδομένων με κατηγορικές ιδιότητες με τη χρήση εύθραυστων υδατογραφημάτων. Στον αλγόριθμό που προτείνεται, γίνεται διαχωρισμός των πλειάδων σε ομάδες, στη συνέχεια τα υδατογραφήματα ενσωματώνονται και επαληθεύονται σε κάθε ομάδα ανεξάρτητα. Η υδατογράφηση αυτή δεν παρεμβαίνει καθόλου στα δεδομένα. Οποιεσδήποτε τροποποιήσεις μπορούν να εντοπιστούν και να απομονωθούν σε επίπεδο ομάδας. Η θεωρητική ανάλυση δείχνει ότι η αποτελεσματικότητα της ανίχνευσης είναι πολύ υψηλή.

Οι δημοσιεύσεις [61-65] επικεντρώνονται στα εύθραστα σχήματα υδατογράφησης που στοχεύουν στον έλεγχο της ακεραιότητας. Το άρθρο [61] προτείνει ένα εύθραυστο σχήμα υδατοσήμανσης για τον εντοπισμό κακόβουλων τροποποιήσεων σε δεδομένα. Με τον προτεινόμενο αλγόριθμο, οι πλειάδες χωρίζονται αρχικά σε ομάδες και στη συνέχεια γίνεται η ενσωμάτωση και η επαλήθευση σε κάθε ομάδα ξεχωριστά. Οι τροποποιήσεις μπορούν να ανιχνευθούν αλλά και να εντοπιστούν σε επίπεδο ομάδας. Το εύθραυστο σχήμα του άρθρου [62], βασίζεται στην προσέγγιση μηδενικής υδατοσήμανσης όπου το υδατογράφημα δημιουργείται από το περιεχόμενα των αρχικών δεδομένων αλλά δεν εισάγει παραμορφώσεις στα δεδομένα.

Το άρθρο [63] προτείνει ένα αλγόριθμο χωρίς παραμόρφωση που δεν στρεβλώνει τα αρχικά δεδομένα. Επιπρόσθετα η παραποιημένη περιοχή μπορεί να εντοπιστεί σε επίπεδο ομάδας. Η εργασία [64] προτείνει ένα ημι-εύθραυστο σχήμα υδατογραφήματος για επαλήθευση της ακεραιότητας που πέρα από τον εντοπισμό τροποποιήσεων των δεδομένων, το σχήμα υποστηρίζει τις περιοδικές ενημερώσεις, χωρίς εκ νέου υδατοσήμανση. Η ενσωμάτωση του υδατογραφήματος περιλαμβάνει την μετατροπή μικρών χαρακτήρων σε κεφαλαία και το αντίστροφο με τη διατήρηση της σημασίας των κειμένων. Η ενσωμάτωση σε επίπεδο ομάδων διευκολύνει τον

μερικό εντοπισμό της παραβίασης. Το άρθρο [65] αναλύει την υδατογράφηση χωρίς τροποποιήσεις. Η μέθοδος που υιοθετεί, χωρίζει τα δεδομένα σε ομάδες τετραγωνικών πινάκων. Στη συνέχεια, τα υδατογραφήματα για κάθε ομάδα, καταχωρούνται σε μια τρίτη αρχή πιστοποίησης. Η επαλήθευση πραγματοποιείται με τον υπολογισμό ορίζουσας πίνακα και ελάσσονος διαγώνιας για κάθε ομάδα. Ως αποτέλεσμα, η παραβίαση μπορεί να ανιχνευτεί σε επίπεδο ομάδας.

Η ομάδα εργασιών [66, 68, 73] ασχολείται με την προσθήκη αποτυπωμάτων με στόχο την ιχνηλάτιση των αντιγράφων και των πειρατών, ενσωματώνοντας ένα μοναδικό σήμα σε κάθε αντίγραφο που διαμοιράζεται. Στην [66] παρουσιάζεται μια τεχνική για την προσθήκη αποτυπωμάτων σε δεδομένα, επεκτείνοντας το δημοφιλές σχήμα ισχυρής υδατογράφησης των Agrawal-Kiernan(2003) [21]. Η νέα δυνατότητα που παρέχεται ενσωματώνει και ανιχνεύσει μια αυθαίρετη συμβολοσειρά από bits ως αποτύπωμα. Τα πειράματα έδειξαν την ανθεκτικότητα των αποτυπωμάτων σε διαφορετικές επιθέσεις. Το σχήμα είναι συμμετρικό αφού ο παραγωγός και ο καταναλωτής αγοραστής, διαθέτουν το ίδιο επισημασμένο αντίγραφο. Μια διαφορετική παραλλαγή του αλγόριθμου Agrawal-Kiernan(2003) [73], παρεμβαίνει σε τυχαίες πλειάδες, χαρακτηριστικά και bits μικρότερης σημασίας. Η εργασία προσθέτει βαρύτητες στα χαρακτηριστικά και αλλάζει την πιθανότητα επιλογής τους για μαρκάρισμα. Τα πειράματα δείχνουν αυξημένη ανθεκτικότητα σε επιθέσεις.

Η δημοσίευση [67] αναλύει τις ευπάθειες αλγορίθμων προσθήκης αποτυπωμάτων σε δεδομένα σε επιθέσεις συσχέτισης. Αναπτύσσει τεχνικές μετριασμού που μπορούν να βελτιώσουν την ανθεκτικότητα των σημερινών σχημάτων. Το άρθρο [68] επεξηγεί ότι η υδατογράφηση και τα αποτυπώματα είναι δύο διαδικασίες που αποδεικνύουν την ιδιοκτησία των δεδομένων και τη νόμιμη κατοχή τους. Για την κοινή χρήση δεδομένων από ένα πάροχο, ο ιδιοκτήτης ενσωματώνει ένα αποτύπωμα, μια συμβολοσειρά που σχετίζεται με τον τον παραλήπτη. Η προτεινόμενη διαδικασία αξιοποιεί την τυχαιότητα για να εγγυηθεί την ιδιωτικότητα και την προσθήκη αποτυπώματος σε δεδομένα. Οι ίδιοι συγγραφείς σε άλλη δημοσίευση [69] επισημαίνουν την ευπάθεια των σχημάτων υδατογράφησης σε επιθέσεις συσχέτισης που μπορεί να είναι κατά χαρακτηριστικό ή κατά πλειάδα. Η αντιμετώπιση του ζητήματος προτείνει ένα ισχυρό ψηφιακό αποτύπωμα στη φάση της μετα-επεξεργασίας. Τα πειράματα επιθέσεων συσχέτισης έγιναν σε μεγάλες και πραγματικές βάσεις δεδομένων. Η εργασία [70] ασχολείται με την απόδειξη ιδιοκτησίας και τη μη εξουσιοδοτημένη διανομή δεδομένων, προτείνοντας πιο γενικές λύσεις για ΜΗ αριθμητικά ή ευαίσθητα σε σφάλματα δεδομένα. Η πρώτη τεχνική χρησιμοποιεί μετάθεση ορισμένων τιμών χαρακτηριστικών μεταξύ προεπιλεγμένων πλειάδων. Η δεύτερη τεχνική, προσθέτει πλειάδες ελέγχου σε αντίθεση με άλλες τεχνικές που επηρεάζουν τα υπάρχοντα δεδομένα με θόρυβο. Το άρθρο [71] προτείνει μια αρχιτεκτονική αποτυπωμάτων με το όνομα «NoFiA» για την αποτροπή παράνομων διανομών. Το σχήμα αυτό ενσωματώσει μοναδικά αποτυπώματα για κάθε χρήστη χωρίς να τροποποιεί τα δεδομένα. Στην αρχιτεκτονική αυτή ελεγχόμενης διάθεσης δεδομένων, τα δεδομένα και οι παράμετροι διατηρούνται ξεχωριστά, ενώ το αποτύπωμα δημιουργείται κατά την πρόσβαση στα δεδομένα μέσω ερωτημάτων.

Η δημοσίευση [72] ασχολείται με την αποφυγή "collusions"-συμπαιγνίας. Ο όρος περιγράφει το συνδυασμό πολλών αντιγράφων στην παραγωγή ενός νέου αντιγράφου ως αντίμετρο στα ψηφιακά αποτυπώματα. Η εργασία παρουσιάζει ένα μηχανισμό εισαγωγής αποτυπωμάτων που αποφεύγει την επίθεση συμπαιγνίας σε αριθμητικά δεδομένα. Το σύστημα διατηρεί την ποιότητα των δεδομένων κρατώντας σταθερές στατιστικές παραμέτρους (μέση τιμή, διακύμανση, τυπική απόκλιση). Μια συνάρτηση κατακερματισμού χρησιμοποιείται στη δημιουργία μοναδικής ακολουθίας χαρακτηριστικών για επισήμανση που δημιουργεί ένα τυχαίο

και μοναδικό μοτίβο εισαγωγής για κάθε καταναλωτή. Τα μοτίβα αυτά δεν μπορούν να συνδυαστούν για επιθέσεις συμπαιγνίας.

Η εργασία [74] δεν αφορά βάσεις δεδομένων αλλά περιλαμβάνει μια εκτενή ιστορική αναδρομή της χρήσης των βιομετρικών αποτυπωμάτων και των επιθέσεων σε αυτά. Περιλαμβάνει την αρχιτεκτονική ενός συστήματος βιομετρίας που ενσωματώνει τα στάδια τη εγγραφής, της πιστοποίησης και της ταυτοποίησης, μιας αρχιτεκτονικής που θα μπορούσε να χρησιμοποιηθεί κατά αναλογία και σε σύνολα δεδομένων.

Το άρθρο [75] αναφέρεται πάλι σε βιομετρικά αποτυπώματα και προτείνει ένα πλαίσιο προσθήκης αποτυπωμάτων σε βάσεις δεδομένων που χρησιμοποιεί σειρές bits από βιομετρικές ψηφιακές υπογραφές. Η πειραματική μελέτη απέδειξε ότι η προτεινόμενη τεχνική ειδικά όταν συνδυαστεί και με μηχανική μάθηση, είναι ανθεκτική απέναντι σε επιθέσεις επιλογής υποσυνόλου, συνδυασμού και αντιστοίχισης, συμπαιγνίας, διαγραφής, εισαγωγής και αλλοίωσης.

Η έρευνα [76] αφορά τη συγκριτική αξιολόγηση επτά διαφορετικών προσεγγίσεων ισχυρής υδατοσήμανσης με βάση προκαθορισμένα κριτήρια. Η εργασία αναφέρει την ανάγκη διαθεσιμότητας ανοιχτών δεδομένων για πειράματα καθώς και μια προσέγγιση αξιολογήσεων που μετρά την αποτελεσματικότητα και αποδοτικότητα των σχημάτων με βάση κοινά αποδεκτές μετρικές και δείκτες.

Η τεχνική της αναστρέψιμης υδατοσήμανσης, υποστηρίζει την ανάκτηση των δεδομένων από την υδατογραφημένη μορφή τους, ενώ τα ΜΗ αναστρέψιμα σχήματα επικεντρώνονται μόνο στην προστασία των δικαιωμάτων ιδιοκτησίας. Αυτή η ανάγκη προέκυψε από δεδομένα που δεν μπορούν να αλλοιωθούν καθόλου όπως τα ιατρικά και τα γενετικά δεδομένα, οι τραπεζικοί λογαριασμοί, οι αριθμοί καρτών, οι αριθμοί ταυτοποίησης προσώπων (ΑΦΜ, ΑΜΚΑ, διαβατήριο, ταυτότητα). Η έρευνα [86] παρουσιάζει μια επισκόπηση όλων των αναστρέψιμων μεθόδων υδατογράφησης.

Το άρθρο [87] παρουσιάζει ένα ισχυρό σχήμα υδατογράφησης που χρησιμοποιεί μια $3^{ου}$ βαθμού καμπύλη Bezier (ως υδατόσημο), υπολογίζει συντεταγμένες x(t),y(t) της καμπύλης με βάση τιμές hash που συνδυάζουν το πρωτεύον κλειδί κάθε πλειάδας με ένα μυστικό κλειδί και ενημερώνει δεκαδικά ψηφία των αριθμητικών πεδίων. Ο αλγόριθμος σε πειράματα διαγραφής και μεταβολής υποσυνόλου, έδειξε ανθεκτικότητα σε υψηλά ποσοστά επίθεσης (>80%).

Η εργασία [88], εξηγεί μια μέθοδο εισαγωγής υδατοσήμανσης σε δεδομένα που χρησιμοποιεί ελλειπτικές κρυπτογραφικές καμπύλες. Ο αλγόριθμος υπολογίζει σημεία της καμπύλης και μαρκάρει δεκαδικά ψηφία αριθμητικά πεδίων με βάση τις συντεταγμένες τους. Ο αλγόριθμος εξετάστηκε σε απειλές διαγραφής, μεταβολής και εισαγωγής υποσυνόλου και τα αποτελέσματα παρουσιάζουν υψηλή ανθεκτικότητας σε υψηλά ποσοστά επίθεσης (>70%).

Το άρθρο [89] εισάγει τον όρο της "μόνιμης υδατοσήμανσης" ως μέσο ελέγχου ακεραιότητας και αναγνώρισης του ιδιοκτήτη, επιτρέποντας την επιβεβαίωση της βάσης σε πραγματικό χρόνο κατά την εφαρμογή ερωτημάτων ενημέρωσης (εισαγωγής, μεταβολής, διαγραφής). Το υδατόσημο χαρακτηρίζεται ως μόνιμο εφόσον μπορεί να ανιχνευθεί σε κάθε αλλαγή της κατάστασης της βάσης που προκύπτει από τα ερωτήματα.

Τα περισσότερα σχήματα υδατοσήμανσης δεδομένων βασίζονται σε μυστικά κλειδιά. Η ιδιοκτησία αφού αποδειχθεί στο δικαστήριο, αποκαλύπτει το κλειδί που μπορεί να χρησιμοποιηθεί κακόβουλα στο μέλλον. Η εργασία [90] προτείνει ένα σχήμα το οποίο ανιχνεύει το υδατόσημο με τη χρήση δημόσιου κλειδιού ενώ δεν αλλοιώνει τα δεδομένα. Το δημόσιο κλειδί αυτό δεν συνδέεται με κάποιο ιδιωτικό κλειδί όπως γίνεται συνήθως στην ασύμμετρη κρυπτογράφηση, αλλά συνδέεται με το ID του χρήστη μέσα από πιστοποιητικά υδατογράφησης που διατηρούνται σε αρχή πιστοποίησης. Το σχήμα δημιουργεί ένα δημόσιο υδατόσημο από συλλογή των MSBs

χαρακτηριστικών. Το σχήμα είναι αποδοτικό στις ενημερώσεις της βάσης και έχει δοκιμασθεί σε όλες τις γνωστές μορφές επιθέσεων.

Η δημοσίευση [91] προτείνει ένα ισχυρό και αναστρέψιμο σχήμα που προσθέτει ένα στάδιο προεπεξεργασίας πριν τη εισαγωγή και ένα ακόμα στάδιο ανάκτησης των αρχικών δεδομένων στο τέλος. Η προεπεξεργασία αφορά την επιλογή των χαρακτηριστικών με στόχο την ελάχιστη αλλοίωση της χρησιμότητας των δεδομένων χρησιμοποιώντας γενετικούς αλγορίθμους.

Το σχήμα της εργασίας [92] εισάγει ένα τυφλό ισχυρό και αναστρέψιμο αλγόριθμο με όνομα "πρόβλεψη σφάλματος σε ακέραιους αριθμούς" που μπορεί να εφαρμοσθεί σε ιατρικά ή στρατιωτικά δεδομένα που δεν επιδέχονται αλλοίωση. Το σχήμα προτείνει ένα πρωτόκολλο χρονοσήμανσης μέσα από μια αρχή πιστοποίησης. Η αλλοίωση των δεδομένων είναι αμελητέα ενώ τα πειράματα έδειξαν οτι η ανίχνευση πετυχαίνει ακόμα και σε διαγραφή του 95% των πλειάδων. Ο πίνακας 2 κατηγοριοποιεί την παραπάνω βιβλιογραφία με βάση τον τρόπο της υδατογράφησης και το είδος των δεδομένων.

ΑΑ	**Περιγραφή**	**Πηγές**
1	Επισκόπηση και ταξινόμηση σχημάτων υδατοσήμανσης	26, 28, 36, 41, 47, 52, 74, 76, 93
2	Ανθεκτικά σχήματα υδατοσήμανσης	21, 22, 23, 25, 29, 30, 31, 32, 34, 40, 44, 46, 48, 49, 50, 73, 68, 69,87
3	Εύθραστα σχήματα υδατοσήμανσης	58, 61, 62, 63, 64, 65, 89
4	Ψηφιακά (δακτυλικά) αποτυπώματα	23, 53, 66, 70, 71
5	Υδατοσήμανση ΜΗ αριθμητικών δεδομένων	42,43
6	Υδατοσήμανση ανωνυμοποιημένων δεδομένων	38, 39
7	Αναστρέψιμα σχήματα	27, 86, 91, 92
8	Χρήση αρχής πιστοποίησης	43, 65, 90
9	Εργαλεία αξιολόγησης αλγορίθμων (benchmarking)	24
10	Εισαγωγή ειδικών αποτυπωμάτων	20, 27, 35, 75, 87, 88, 90

Πίνακας 2: Κατηγοριοποίηση της βιβλιογραφίας στην υδατογράφηση δεδομένων

Τέλος, η δημοσίευση [93] συγκεντρώνει όλες τις γνωστές μεθόδους υδατογράφησης σε δεδομένα. Μια διαφορετική ταξινόμηση των μεθόδων φαίνεται στον πίνακα 3.

ΑΑ	**Μέθοδος**	**Σκοποί**	**Δεδομένα**	**Τύπος**
1	Αντικατάσταση LSB	Αναγνώριση, Ιχνηλάτηση	Αριθμητικά	Ανθεκτικό
2	Ενσωμάτωση εικόνας/ήχου	Αναγνώριση	Αριθμητικά	Ανθεκτικό
3	Εισαγωγή πλειάδων	Αναγνώριση	Όλα	Ανθεκτικό
4	Μεταβολή στατιστικών	Αναγνώριση	Κείμενο Κατηγορίες	Ανθεκτικό
5	Αντικατάσταση χαρακτήρων	Αναγνώριση	Κείμενο	Ανθεκτικό
6	Εισαγωγή χαρακτηριστικών	Ακεραιότητα	Όλα	Εύθραυστο
7	Ταξινόμηση πλειάδων	Ακεραιότητα	Όλα	Εύθραυστο

Πίνακας 3:Ταξινόμηση μεθόδων υδατογράφησης

Η επισκόπηση της βιβλιογραφίας αναγνώρισε κάποια κενά στην τεκμηρίωση των υλοποιήσεων.

- Ένας αριθμός δημοσιεύσεων [22, 25, 32, 40, 45, 46, 58, 61, 66] δεν περιγράφει τα δεδομένα που χρησιμοποιήθηκαν στα πειράματα.
- Πολλές δημοσιεύσεις αναφέρουν πειράματα με ένα μόνο dataset [20, 21, 27, 30, 31, 38, 39, 44, 50, 53, 62-72, 75].
- Κάποιες εργασίες υδατογράφησης [21, 24] παρουσιάζουν πειράματα σε datasets ακέραιων αριθμών (χωρίς δεκαδικά ψηφία).
- Κάποιες εργασίες [24, 29, 33, 35, 36, 48] αναφέρουν πειραματισμό με 3-4 διαφορετικά datasets.
- Οι περισσότερες εργασίες αναφέρουν ως προαπαιτούμενο κάποια βάση δεδομένων στην οποία εισάγουν τα datasets για επεξεργασία.
- Τα γραφήματα των αποτελεσμάτων εργασιών δεν καλύπτουν γενικά όλες τις μορφές των γνωστών επιθέσεων αλλά επικεντρώνονται σε κάποιες.

Η παρούσα εργασία με στόχο την πληρότητα, τη διάθεση των εργαλείων, των πειραμάτων και των αποτελεσμάτων στην επιστημονική κοινότητα προσπάθησε να καλύψει τα παραπάνω κενά μέσα από πειράματα με πολλαπλά datasets, με μίγμα ακεραίων και δεκαδικών αριθμητικών τιμών, φροντίζοντας σε μια πλήρη τεκμηρίωση της υλοποίησης, των σεναρίων και της οπτικοποίησης των αποτελεσμάτων.

Οι εργασίες που περιέχουν επισκόπηση και ταξινόμηση των γνωστών προσεγγίσεων είναι πολύ αναλυτικές [26, 28, 36, 41, 47, 52, 74, 76] όμως αποτελούν στιγμιότυπα της χρονικής στιγμής που γίνεται η έρευνα και έχουν στατικό χαρακτήρα. Θα ήταν πολύ χρήσιμο εάν υπήρχε ένα κοινό αποθετήριο στο οποίο θα καταγραφόταν κάθε υλοποίηση υδατογράφησης με διάθεση των υλοποιήσεων σε μορφή ανοιχτού κώδικα και των datasets. Επίσης τα πειράματα και τα αποτελέσματα θα ήταν πολύ χρήσιμα σε επεξεργάσιμη και μηχαναγνώσιμη μορφή για την εύκολη σύγκριση τους. Τα θέματα της απόδοσης των υλοποιήσεων έχουν μελετηθεί σε ένα βαθμό, λείπει όμως μια συνολική ποσοτική και ποιοτική σύγκριση όλων των προσεγγίσεων καθώς είναι περιορισμένα και τα εργαλεία για το σκοπό αυτό.

Η έρευνα στην υδατογράφηση ΒΔ περιλαμβάνει περιοχές που φαίνεται να μην έχουν εξερευνηθεί αρκετά, όπως είναι η υδατογράφηση κειμένου με βάση τη σημασιολογία, η ανάπτυξη διεθνών προτύπων που θα ενσωματώνουν την ταυτότητα του δημιουργού στο περιεχόμενο, η καταχώρηση των δημιουργών-έργων σε παγκόσμια βάση δεδομένων, η χρήση blockchain στην άρρηκτη σύνδεση των έργων με τον δημιουργό, η ανάπτυξη ενός μοντέλου αυτοματοποιημένης αναγνώρισης και οικονομικής αποζημίωσης των δημιουργών.

3 Πνευματική ιδιοκτησία ψηφιακών αγαθών, νομικά και τεχνολογικά μέσα για την προστασία τους

Η πνευματική ιδιοκτησία, αφορά στην ηθική αναγνώριση του δημιουργού και στα προνόμια που απολαμβάνει στη αποκλειστική χρήση και οικονομική εκμετάλλευση των πνευματικών έργων του. Η αναγνώριση των παραπάνω δικαιωμάτων, αποτελεί μορφή ανταμοιβής και κινήτρων που προσφέρει μια οργανωμένη κοινωνία στη δημιουργική και παραγωγική δραστηριότητα των πνευματικών ανθρώπων που επενδύουν σε χρόνο, εργασία και οικονομικούς πόρους στη δημιουργία καινοτόμων έργων [1]. Ένα έργο για να μπορεί να γίνει αντικείμενο πνευματικής ιδιοκτησίας, πρέπει να είναι πνευματικής φύσης, να είναι καινοτόμο και να υπάγεται στο χώρο της τέχνης, της επιστήμης ή της τεχνολογίας, προσφέροντας καινούργια στοιχεία γνώσης και πολιτισμού που συμβάλλουν στην ανταγωνιστικότητα και στην οικονομική ανάπτυξη. Η πρωτοτυπία ενσωματώνει την συνεισφορά του δημιουργού, τη συμβολή του στην τέχνη ή την επιστήμη και δεν αναφέρεται τόσο στην αξία ή στο μέγεθος της προσπάθειας αλλά στη μοναδικότητα του αποτελέσματος που παράγεται.

Οι βάσεις, οι συλλογές και τα σύνολα δεδομένων αποτελούν περίπτωση συλλεκτικού έργου εφόσον έχουν δημιουργηθεί με συστηματικό τρόπο και μπορεί να περιλαμβάνουν έργα φυσικά, καλλιτεχνικά, λογοτεχνικά, ψηφιακά πολυμεσικά (κείμενο, ήχος, εικόνα, βίντεο), αναλυτικά, στατιστικά δεδομένα ή διαγράμματα που έχουν συλλεχθεί από έρευνες ή μελέτες. Για τα ελληνικά δεδομένα, το αρχείο της ΕΡΤ θεωρείται μια μεγάλη συλλογή δεδομένων στο χώρο του πολιτισμού. Οι σύγχρονες συλλογές δεδομένων είναι συνήθως ψηφιακές, αποθηκευμένες σε ψηφιακά μέσα και αποτελούν πνευματικό δημιούργημα όχι λόγω της πρωτοτυπίας του περιεχομένου αλλά λόγω της συλλογής και της οργάνωσής του.

Η προστασία της πνευματικής ιδιοκτησίας μια συλλογής, αναφέρεται στη νομική προστασία του δημιουργού της συλλογής (όχι στους δημιουργούς των επιμέρους στοιχείων) και δίνει αποκλειστικά δικαιώματα στην επεξεργασία, στη μετατροπή, στην αναπαραγωγή ή στη διανομή της συλλογής με διάφορα μέσα και σε διάφορες μορφές στο ευρύ κοινό [56].

Ο κατασκευαστής μιας βάσης δεδομένων ή συλλογής, ορίζεται ως το φυσικό ή νομικό πρόσωπο που επενδύει στη δημιουργία της ενώ αποκτά το αποκλειστικό δικαίωμα της αξιοποίησης του συνόλου ή μεγάλου μέρους της εφόσον, η απόκτηση του περιεχομένου είναι αποτέλεσμα οικονομικής επένδυσης. Ο κατασκευαστής αποκτά το αποκλειστικό δικαίωμα της αναζήτησης και εξαγωγής δεδομένων από τη βάση του και προστατεύεται για δεκαπέντε χρόνια μετά τη δημιουργία της βάσης [56]. Ο κατασκευαστής-δικαιούχος δύναται να παραχωρήσει δικαιώματα χρήσης της βάσης σε νόμιμους χρήστες οι οποίοι μπορούν να αναζητήσουν ή να αντλήσουν υλικό από αυτήν. Οι νόμιμοι χρήστες δεν αποκτούν απεριόριστα δικαιώματα σε μια βάση. Μπορούν να προβαίνουν σε αναπαραγωγή αλλά όχι σε επαναχρησιμοποίηση ουσιώδους μέρους της βάσης δεδομένων, μόνο όταν πρόκειται για εκπαιδευτικούς ή ερευνητικούς σκοπούς. Όταν ο σκοπός εξυπηρετεί τη δημόσια ασφάλεια ή κάποια δικαστικής διαδικασία, μπορεί εκτός της αναπαραγωγής να γίνει και επαναχρησιμοποίηση ουσιώδους μέρους της βάσης [56].

Παράλληλα με τη νομική προστασία των βάσεων δεδομένων υπάρχουν και τεχνολογικά περιοριστικά μέτρα που θέτουν κανόνες στην πρόσβαση και τη χρήση των δεδομένων. Τα τεχνολογικά μέτρα, αξιοποιούνται σήμερα από τους δικαιούχους για την αποτροπή των παράνομων προσβάσεων και χρήσεων των έργων τους και περιλαμβάνουν την κρυπτογραφία, την αναγνωριστική ετικέτα, το υδατογράφημα και τα ολοκληρωμένα συστήματα διαχείρισης δικαιωμάτων και αντιγράφων. Τα τεχνολογικά μέτρα με τη σειρά τους υπόκεινται επίσης σε νομική προστασία που προστατεύει την εξουδετέρωσή τους μέσω τεχνολογικών αντίμετρων και επιβάλει αστικές και ποινικές κυρώσεις.

3.1 Τα νομικά μέσα προστασίας

Η κοινωνία της πληροφορίας έχει διεθνή διάσταση [1]. Η προστασία της πνευματικής ιδιοκτησίας ξεκινά το 1886 με τη Διεθνή Σύμβαση της Βέρνης, η οποία τεκμηριώνει την αρχή της εξομοίωσης των ξένων δημιουργών ή έργων προς τους υπηκόους της χώρας όπου ζητείται η προστασία τους [1,5]. Η τελευταία αναθεώρηση αυτής της σύμβασης έγινε στο Παρίσι το 1971 και κυρώθηκε από την Ελλάδα με το Ν.100/1975. Η αναγνώριση των συγγενικών δικαιωμάτων σε διεθνές επίπεδο έγινε με τη Διεθνή Σύμβαση της Ρώμης το 1961 που προστατεύει τους ερμηνευτές ή εκτελεστές καλλιτέχνες, τους παραγωγούς φωνογραφημάτων και τους ραδιοτηλεοπτικούς οργανισμούς. Η Ελλάδα την επικύρωσε με το Ν.2054/1992 (ΦΕΚ Α΄ 104).

Η Συμφωνία TRIPS για τα Δικαιώματα Διανοητικής Ιδιοκτησίας στον τομέα του Εμπορίου, περιλαμβάνεται στην Τελική Πράξη του Γύρου της Ουρουγουάης (Μαράκες 1994) που κυρώθηκε από την Ελλάδα με το Ν.2290/1995 (ΦΕΚ Α΄28). Η Συμφωνία TRIPS άρχισε να δεσμεύει τα περισσότερα κράτη την 1η Ιανουαρίου 1996 και έχει ευρύ πεδίο εφαρμογής γιατί καλύπτει τόσο τη βιομηχανική όσο και την πνευματική ιδιοκτησία και τα συγγενικά δικαιώματα που βρίσκονται στο επίκεντρο των διεθνών οικονομικών σχέσεων στο πεδίο δράσης του Παγκόσμιου Οργανισμού Εμπορίου (ΠΟΕ). Βασικός σκοπός της Συμφωνίας TRIPS είναι ο

περιορισμός των στρεβλώσεων και των εμποδίων για το διεθνές εμπόριο, η αποτελεσματική και η επαρκής προστασία των δικαιωμάτων πνευματικής και βιομηχανικής ιδιοκτησίας καθώς και η διασφάλιση των μέτρων και των διαδικασιών για την επιβολή των δικαιωμάτων αυτών.

Στο επίπεδο της Ευρωπαϊκής ένωσης, έχουν εκδοθεί έξι οδηγίες που καλύπτουν το δημιουργό και τα συγγενικά δικαιώματα [1,2,4,5].

i. Η Οδηγία 91/250/ΕΟΚ για τη νομική προστασία του λογισμικού που υπολογίζεται ως λογοτεχνικό έργο στο πλαίσιο της πνευματικής ιδιοκτησίας.
ii. Η Οδηγία 92/100/ΕΟΚ που καλύπτει το δικαίωμα εκμίσθωσης, δανεισμού και εναρμονίζει τα συγγενικά δικαιώματα.
iii. Η Οδηγία 93/83/ΕΟΚ αφορά τα πνευματικά δικαίωμα στις ραδιοτηλεοπτικές καλωδιακές και δορυφορικές μεταδόσεις στην Ευρώπη.
iv. Η Οδηγία 93/98/ΕΟΚ εναρμονίζει τις χρονικές διάρκειες προστασίας πνευματικής ιδιοκτησίας και συγγενών δικαιωμάτων στα εβδομήντα έτη μετά το θάνατο του δημιουργού και στα πενήντα έτη για τα συγγενικά δικαιώματα.
v. Η Οδηγία 96/9/ΕΟΚ αφορά τη νομική προστασία των βάσεων δεδομένων που έχουν ψηφιακή μορφή ή είναι παραδοσιακή συλλογή δεδομένων. Η πνευματική ιδιοκτησία αφορά τις βάσεις δεδομένων που εμφανίζουν πρωτοτυπία ενώ το οικονομικό δικαίωμα εκμετάλλευσης, προστατεύει την επένδυση του κατασκευαστή μιας βάσης δεδομένων, απαγορεύοντας την εξαγωγή ή την επαναχρησιμοποίηση μέρους ή του συνόλου της βάσης.
vi. Η Οδηγία 2001/29 προστατεύει νομικά το δημιουργό και αντιστοιχεί στις διατάξεις των συνθηκών για το διαδίκτυο, για την πνευματική ιδιοκτησία, τις εκτελέσεις και τα φωνογραφήματα. Η οδηγία αυτή ρυθμίζει τα ψηφιακά δικαιώματα αναπαραγωγής, παρουσίασης και διανομής στο κοινό, επιτρέποντας στο δικαιούχο την ιχνηλάτηση της διαδρομή των έργων του στο Διαδίκτυο.

Η πνευματική ιδιοκτησία και οι δημιουργοί των πρωτότυπων έργων, προστατεύονται με νομικά και τεχνολογικά μέσα και στην Ελλάδα. Η πνευματική ιδιοκτησία θεμελιώνεται συνταγματικά [54] αφού σχετίζεται με την ελεύθερη ανάπτυξη της προσωπικότητας (Αρ. 5), την ελευθερία της έκφρασης και του τύπου (Αρ. 14), την ελευθερία της τέχνης, της επιστήμης, της έρευνας και της διδασκαλίας (Αρ.16), την προστασία της υλικής και της άυλης ιδιοκτησίας (Αρ. 17), την οικονομική ελευθερία του ατόμου στην αξιοποίηση των πνευματικών έργων του (Αρ. 5).

Η ελληνική νομοθεσία υιοθετεί και προσαρμόζει στο εθνικό πλαίσιο τις ρυθμίσεις της Ευρωπαϊκής Οδηγίας 2001/29 με το άρθρο 81 Ν.3057/2002 (ΦΕΚ Α΄ 239/10-10-2002), ενώ κυρώνει τις δύο Συνθήκες του διαδικτύου με τους Ν. 3183/2003 και Ν.3184/2003. Η ελληνική νομοθεσία εναρμονίζεται με το ευρωπαϊκό σύστημα πνευματικής ιδιοκτησίας και υιοθετεί τις παραδοσιακές αρχές της θεωρία του σκοπού, του έγγραφου τύπου, της ποσοστιαίας αμοιβής, της προστασία του ηθικού δικαιώματος και της «αρχής της αλήθειας» όπου ο αρχικός δικαιούχος είναι ο φυσικός δημιουργός και ισχύει και στην περίπτωση δημιουργίας έργου κατά παραγγελία. Η ελληνική νομοθεσία αναγνωρίζει το λογισμικό και τις βάσεις δεδομένων και τα προστατεύει ως έργα λόγου, ενώ η προσβολή των δικαιωμάτων προβλέπει αυστηρές ποινικές κυρώσεις.

Η πνευματική ιδιοκτησία, η ελευθερία τη δημιουργικότητας και η οικονομική ελευθερία της αξιοποίησης των υλικών και των άυλων αγαθών αναφέρονται και στο Χάρτη Θεμελιωδών Δικαιωμάτων της Ευρώπης [55]. Συγκεκριμένα, αναφέρονται η ελευθερία της τέχνης και της επιστήμης (Αρ. 13), η επιχειρηματική και η οικονομική ελευθερία (Αρ. 16) καθώς και η διανοητική ιδιοκτησία (Αρ. 17).

3.2 Τα τεχνολογικά μέσα προστασίας

Οι τεχνολογικές προσεγγίσεις στην προστασία των πνευματικών δικαιωμάτων, περιλαμβάνουν ένα μίγμα μεθόδων :

1. Την υδατοσήμανση που αφορά το αντικείμενο της παρούσης εργασίας και αντιμετωπίζει την προστασία του Copyright, την πιστοποίηση της αυθεντικότητας του ψηφιακού περιεχομένου και την αναγνώριση του κατά τη διανομή και την αντιγραφή μέσω παγκόσμιων δικτύων. Εφαρμόζεται σε κείμενο, σε πολυμεσικό περιεχόμενο (εικόνες, video, ηχογραφήσεις) και σε σύνολα δεδομένων.
2. Τα συστήματα αναγνώρισης που αποσκοπούν στην αναγνώριση και στην ανταλλαγή πληροφοριών σχετικών με το δημιουργό.
3. Τα μεταδεδομένα ως πρόσθετες πληροφορίες που συνοδεύουν το περιεχόμενο και βοηθούν στη διαχείριση του Copyright.
4. Τα συστήματα διαχείρισης ψηφιακών δικαιωμάτων (DRM)
5. Τις γλώσσες προγραμματισμού που βασίζονται στα σύνολα μετα-δεδομένων και επιτυγχάνουν τη δια-λειτουργικότητα των συστημάτων διαχείρισης των πνευματικών δικαιωμάτων. Παραδείγματα περιλαμβάνουν την Extensible Access Control Markup Language, την OASIS Rights Language, την Extensible Rights Markup Language, το INDECS Rights Data Dictionary, την Open Digital Rights Language.

Τα λειτουργικά συστήματα μπορούν να ενσωματώσουν λειτουργίες διαχείρισης πνευματικών δικαιωμάτων όπως κάνει η Microsoft με τα Microsoft Windows Rights Management Services. Επίσης κάποιες εφαρμογές μετάδοσης, υποστηρίζουν τη διαχείριση πνευματικών δικαιωμάτων ταυτόχρονα με τη διανομή του περιεχομένου τους. Ενδεικτικά ονόματα είναι η Adobe, το DMD Secure, το IBM EMMS, το Realnetworks, η Sony.

3.3 Η τεχνική της υδατοσήμανσης στα δεδομένα

Η ψηφιακή μορφή έργων τέχνης έγινε πολύ δημοφιλής και ακολούθησε την εξάπλωση του διαδικτύου. Στον ψηφιακό κόσμο, τα ψηφιακά δημιουργήματα μπορούν εύκολα να αντιγραφούν, να επεξεργαστούν, να εκτεθούν και να αναμεταδοθούν σε έντιμους αλλά και σε μη εξουσιοδοτημένους παραλήπτες που τα χρησιμοποιούν παρανόμως για προσωπικό όφελος παραβιάζοντας τα πνευματικά δικαιώματα των δημιουργών.

Η μη εξουσιοδοτημένη χρήση ψηφιακού περιεχομένου περιλαμβάνει :

i. Την παράνομη πρόσβαση σε ψηφιακά έργα στο διαδίκτυο χωρίς την απαραίτητη άδεια.

ii. Την εσκεμμένη φθορά και τροποποίηση του ψηφιακού περιεχομένου, την αλλοίωση της αυθεντικότητας.

iii. Την παραβίαση του δικαιώματος της αναπαραγωγής (copyright) μέσω της εμπορικής διακίνησης χωρίς την απαραίτητη άδεια πώλησης και αναπαραγωγής.

Η ψηφιακή υδατοσήμανση με βάση τη βιβλιογραφία ξεκίνησε τη δεκαετία του ενενήντα ενώ έως σήμερα έχουν προταθεί μέθοδοι υδατογράφησης για όλες τις μορφές ψηφιακών μέσων, τα πολυμέσα, τα ψηφιακά έγγραφα, το λογισμικό και τις βάσεις δεδομένων.

Τα ισχυρά υδατογραφήματα αφορούν την επαλήθευση ιδιοκτησίας ενώ τα εύθραυστα υδατογραφήματα την ανίχνευση της παραβίασης. Το ισχυρό υδατογράφημα παρουσιάζει ανθεκτικότητα σε επιθέσεις και σε νόμιμες τροποποιήσεις δεδομένων των χρηστών με στόχο να καθορίσει την πνευματική ιδιοκτησία. Ο σκοπός ενός εύθραυστου υδατογραφήματος είναι να προσδιορίσει κάθε επίθεση που στρέφεται στην ακεραιότητα του ψηφιακού έργου.

Στις σχεσιακές βάσεις δεδομένων και στα σύνολα δεδομένων, τα δεδομένα οργανώνονται σε μορφή πίνακα που αποτελείται από γραμμές [πλειάδες] και στήλες [γνωρίσματα]. Ένα υδατογράφημα εισάγει με ελεγχόμενο τρόπο ένα σύνολο από μικρά σημάδια στα δεδομένα (marks) σε κάποια γνωρίσματα και σε κάποιες πλειάδες, με μικρή επίδραση στα δεδομένα και με στόχο να καλύπτει πολλές πλειάδες για να είναι ανθεκτικό. Μια επίθεση μπορεί να διαγράψει, να τροποποιήσει ή να αντικαταστήσει μια ή περισσότερες πλειάδες όμως δεν μπορεί να αφαιρέσει στο σύνολο το υδατογράφημα. Η ισχυρή υδατογράφηση δεν προστατεύει την ακεραιότητα των δεδομένων αλλά τη μη εξουσιοδοτημένη αντιγραφή, αποδεικνύοντας την αυθεντικότητα, την πειρατεία και την πλαστογράφηση. Υπάρχουν πολλές υλοποιήσεις υδατογραφημάτων που αναλύονται σε σχετικές δημοσιεύσεις [26, 28, 36, 41, 47, 52, 74, 76].

Από τη φύση του ένα υδατογράφημα τροποποιεί το αντικείμενο που υδατογραφείται εισάγοντας ένα ανεξίτηλο σημάδι στο περιεχόμενο [45] έτσι ώστε η εισαγωγή του σήματος να μην καταστρέφει την αξία του και να είναι δύσκολο για σε έναν τρίτο να αφαιρέσει ή να αλλάξει το σήμα χωρίς να καταστρέψει το περιεχόμενο. Εάν το έργο που πρόκειται να υδατογραφηθεί δεν μπορεί να τροποποιηθεί χωρίς να χαθεί η αξία του τότε δεν μπορεί να εισαχθεί ένα πλαστό υδατογράφημα. Το κρίσιμο ζήτημα είναι να περιοριστούν σε αποδεκτά επίπεδα οι αλλοιώσεις για την προβλεπόμενη χρήση του έργου.

Ένα υδατογράφημα μπορεί να εξυπηρετήσει διαφορετικούς σκοπούς [1] :

i. Προστασία της άδειας αντιγραφής Copyright : Η διαδικασία της ανίχνευσης, παράγει ένα αποτέλεσμα που αποδεικνύει το Copyright του ψηφιακού περιεχομένου με ένα βαθμό βεβαιότητας, που προκύπτει σε σχέση με μια οριοθετημένη τιμή που καθορίζει την ύπαρξη υδατοσήματος. Οι απαιτήσεις της προστασίας του copyright απαιτούν ιδανικά την ενσωμάτωση δύο υδατοσημάτων, ένα για την πιστοποίηση του κατόχου και ένα για την παρεμπόδιση της μη εξουσιοδοτημένης αντιγραφής.

ii. Πιστοποίηση Κατόχου Δικαιωμάτων : Σε μια σύγκρουση του νόμιμου κατόχου με έναν ψεύτικο, ο πραγματικός κάτοχος έχει το κλειδί και μπορεί να αποδείξει ότι είναι ο κάτοχος εκτελώντας το μηχανισμό ανίχνευσης με το κλειδί.

iii. Έλεγχο Αντιγραφής : Ο διαχειριστής του ψηφιακού περιεχομένου εισάγει υδατοσήματα σε κάθε ψηφιακό αντικείμενο που διαθέτει, με τη χρήση ενός σταθερού

κλειδιού. Εξειδικευμένες διεπαφές μπορούν να ανιχνεύσουν το μηχανισμό ανίχνευσης που περιλαμβάνει το εν λόγω κλειδί. Κατά τη χρήση μιας ψηφιακής εικόνας από την εξειδικευμένη διεπαφή, γίνεται ανίχνευση του υδατοσήματος. Κατά την εξαγωγή θετικού αποτελέσματος η διεπαφή εφαρμόζει την εντολή «μην αντιγράφετε» και απαγορεύει την αντιγραφή. Το παράδειγμα αυτό παρουσιάζει την ανάγκη ύπαρξης μιας εξειδικευμένης διεπαφής που ενσωματώνει στις λειτουργίες της μηχανισμούς ανίχνευσης υδατοσήμανσης και περιορισμού λειτουργιών όπως η αντιγραφή ή η εκτύπωση.

iv. Ανίχνευση Εισβολέα : Ο διαχειριστής ενός ψηφιακού περιεχομένου μπορεί να ανακαλύψει αν ένας εισβολέας προσπαθεί να επιτεθεί στο περιεχόμενο αλλάζοντας το. Ο διαχειριστής διατηρεί ένα αρχείο, το οποίο συσχετίζει κλειδιά υδατοσήμανσης με μοναδικά αναγνωριστικά χρηστών στα οποία διατίθεται το περιεχόμενο μετά από αίτημα και με ελεγχόμενο τρόπο. Τα κλειδιά αυτά διανέμονται με μυστικό τρόπο και ο εξουσιοδοτημένος χρήστης διαθέτει ένα μοναδικό ιδιωτικό κλειδί. Όταν ένας εξουσιοδοτημένος χρήστης επιθυμεί να αποθηκεύσει πληροφορία στο ψηφιακό απόθεμα, αποστέλλει την πληροφορία μαζί με το κλειδί υδατοσήμανσης μέσω εξειδικευμένης εφαρμογής που ενσωματώνει το κλειδί στην ψηφιακή εικόνα. Ο διαχειριστής του ψηφιακού περιεχομένου παρατηρεί το αρχείο και ελέγχει τα κλειδιά των εξουσιοδοτημένων χρηστών. Αν το αποτέλεσμα του ελέγχου είναι θετικό τότε το περιεχόμενο χαρακτηρίζεται ως «εγκεκριμένο» και δε διαγράφεται. Σε αντίθετη περίπτωση το περιεχόμενο διαγράφεται. Συνεπώς, μόνο η εξουσιοδοτημένη ομάδα χρηστών επιτρέπεται να αποθηκεύσει πληροφορία στο ψηφιακό περιεχόμενο.

v. Ανίχνευση Συναλλαγών και Μοναδική Αναγνώριση Ψηφιακού Περιεχομένου : Ο διαχειριστής είναι απαραίτητο να διατηρεί ένα αρχείο με αριθμούς και ονόματα. Δεν υπάρχει πλήρης πληροφορία για την ταυτότητα του τελικού χρήστη, αφού αυτός εκτός από εξουσιοδοτημένος χρήστης μπορεί να είναι και κάποιος τρίτος. Αφού το κλειδί για την ανίχνευση είναι μυστικό, απαιτείται η παραγωγή ενός συνδυαστικού κλειδιού που περιλαμβάνει ένα σταθερό κλειδί και μια ακέραια τιμή που παράγεται κατά την εκκίνηση της συναλλαγής και συσχετίζει τον τύπο της συναλλαγής με το πρωτότυπο ψηφιακό περιεχόμενο. Κατά τη διαδικασία της ανίχνευσης, ο ανιχνευτής παίρνει ως αποτέλεσμα έναν κωδικό που οδηγεί σε περισσότερες πληροφορίες σχετικά με το πρωτότυπο περιεχόμενο για τη μοναδική αναγνώριση του ψηφιακού περιεχομένου και της συναλλαγής (τύπος, χρήστης, χρόνος). Αν το ψηφιακό έργο έχει βρεθεί σε κάποιο κόμβο του διαδικτύου, ενώ δεν προβλεπόταν από τη συναλλαγή τότε προκύπτει ζήτημα μη εξουσιοδοτημένης χρήσης και μέσω των πληροφοριών της συναλλαγής αναγνωρίζεται ποιός είχε αρχικά προμηθευτεί το ψηφιακό περιεχόμενο.

Ένα υδατογράφημα ενσωματώνεται στα ψηφιακά δεδομένα που προστατεύει και πρέπει να είναι ανθεκτικό, δηλαδή να μην αλλοιώνεται από επεξεργασία (μελλοντικές συναλλαγές στη βάση δεδομένων) ή από στοχευμένες επιθέσεις πάνω του. Ο Πίνακας 4 συνοψίζει τις βασικές ιδιότητες ενός υδατογραφήματος [1] που αφορούν την ασφάλεια, την ανθεκτικότητα, την αδιορατότητα την ανακτησιμότητα και τη δυνατότητα εξαγωγής του υδατογραφήματος.

ΑΑ	*Ιδιότητα*	*Περιγραφή*
1	*Ασφάλεια*	*Για να διασφαλισθεί η μυστικότητα της ενσωματωμένης πληροφορίας αξιοποιούνται ασφαλή κλειδιά και γεννήτριες ψευδοτυχαίων αριθμών στη διαδικασία της εισαγωγής και της εξαγωγής. Ένα ενσωματωμένο υδατογράφημα είναι μη ανιχνεύσιμο αλλά ακόμα και αν ανιχνευθεί, η επισκόπησή του απαιτεί το σωστό κλειδί. Η παραγωγή, διαχείριση και διανομή των κλειδιών έχουν μεγάλη σημασία όπως και σε κάθε άλλη κρυπτογραφική ενέργεια.*
2	*Ανθεκτικότητα*	*Για να είναι ανθεκτικά τα υδατογραφήματα πρέπει να έχουν όσο το δυνατό μεγαλύτερο μέτρο που να μην αλλοιώνει όμως την αρχική πληροφορία σε μεγάλο βαθμό. Η επεξεργασία δεδομένων περιλαμβάνει κάθε τροποποίηση που μπορούν να υποστούν τα δεδομένα. Μία επίθεση υποδεικνύει την εκμετάλλευση των δεδομένων με σκοπό την αναπαραγωγή, την καταστροφή ή την εξαγωγή των υδατογραφημάτων που εμπεριέχουν. Αν οι αλγόριθμοι ανίχνευσης των υδατογραφημάτων και τα κλειδιά είναι δημόσια, τότε τα υδατογραφήματα είναι ευάλωτα σε επιθέσεις. Συνήθως οι αλγόριθμοι είναι γνωστοί ενώ τα κλειδιά είναι μυστικά όπως συμβαίνει και στην κρυπτογραφία.*
3	*Αδιορατότητα*	*Η διαδικασία ενσωμάτωσης δεδομένων δε θα πρέπει να εισάγει κάποιο ορατό αντικείμενο στα αρχικά δεδομένα. Ιδανικά η ενσωμάτωση ενός υδατογραφήματος πρέπει να βρίσκεται λίγο χαμηλότερα από το όριο της αντίληψης.*
4	*Ανάκτηση*	*Η ανάκτηση ενός υδατογραφήματος είναι πιο εύκολη αν τα αρχικά δεδομένα είναι διαθέσιμα. Οι σύγχρονες τεχνικές υδατοσήμανσης δεν απαιτούν τη διαθεσιμότητα των αρχικών δεδομένων για την ανάκτηση του υδατογραφήματος.*
5	*Εξαγωγή*	*Υπάρχουν δύο κατηγορίες υδατοσήμανσης : αυτή που ενσωματώνει ένα συγκεκριμένο σύνολο από πληροφορίες και μετά το ελέγχει και αυτή που ενσωματώνει αυθαίρετη πληροφορία στα αρχικά δεδομένα. Στον πρώτο τύπο, η εξακρίβωση της ύπαρξης ενός γνωστού υδατογραφήματος, είναι επαρκής στην προστασία της πνευματικής ιδιοκτησίας. Για το δεύτερο τύπο, η ενσωμάτωση αυθαίρετης πληροφορίας είναι χρήσιμη για την ανίχνευση της ενσωματωμένης πληροφορίας.*

Πίνακας 4: Οι ιδιότητες ενός ψηφιακού υδατογραφήματος

Οι τεχνικές που αντιμετωπίζουν την αυθαίρετη χρήση ψηφιακού περιεχομένου βασίζονται στην κρυπτογραφία, στις ψηφιακές υπογραφές, στα υδατογραφήματα και περιλαμβάνουν :

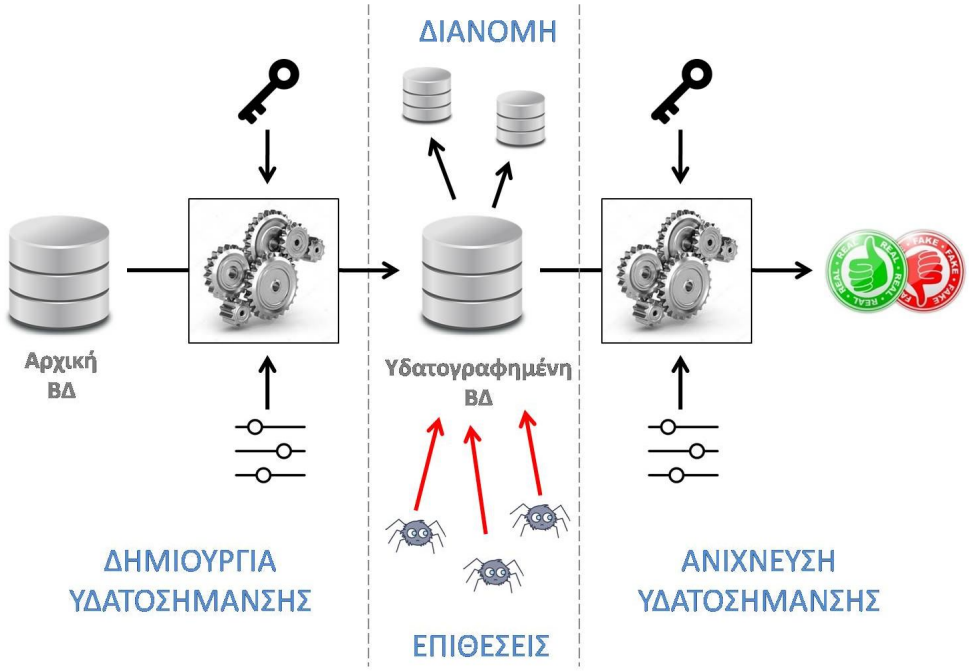
i. Έλεγχο και περιορισμό της πρόσβασης στο αρχείο ή την υπηρεσία (με password)
ii. Υλικό φορέα που δεν μπορεί να αντιγραφεί (πρόσβαση με ειδικό κλειδί)
iii. Περιορισμό της μόνιμης αναπαραγωγής με τη χρήση αύξουσας αρίθμησης και μέγιστου αριθμού αντιγράφων

Η προστασία της εμπιστευτικότητας του περιεχομένου μέσα από ελεγχόμενη πρόσβαση, μπορεί να γίνει με τη χρήση της κρυπτογραφίας. Η συμμετρική κρυπτογραφία επιτρέπει την (από)κρυπτογράφηση περιεχομένου με βάση ένα κοινό κλειδί ενώ η ασύμμετρη κρυπτογραφία (δημόσιου κλειδιού), επιτρέπει την κρυπτογράφηση του περιεχομένου με τη χρήση ενός δημόσιου κλειδιού και την αποκρυπτογράφηση με κάποιο αντίστοιχο ιδιωτικό κλειδί.

Οι διαδικασίες κρυπτογράφησης και αποκρυπτογράφησης βασίζονται σε δημόσιους αλγόριθμους, χρησιμοποιούνται ευρέως σε πραγματικό χρόνο σε όλα τα πρωτόκολλα ασφαλούς επικοινωνίας ενσύρματων, ασύρματων και φορητών δικτύων, στην αποθήκευση δεδομένων ενώ υπάρχουν ειδικοί μηχανισμοί για τη δημιουργία, τη διαχείριση και την ασφαλή διανομή των μυστικών κλειδιών. Οι ψηφιακές υπογραφές βασίζονται επίσης στην κρυπτογράφηση για να ελέγξουν την αυθεντικότητα του αποστολέα. Ο δημιουργός παράγει μία ψηφιακή υπογραφή με τη χρήση ενός ιδιωτικού κλειδιού και ένας δημόσιος αλγόριθμος εξακρίβωσης ελέγχει αν το περιεχόμενο συμβαδίζει με την ψηφιακή υπογραφή.

Η ψηφιακή υδατοσήμανση προέρχεται από αρχαίες τεχνικές απόκρυψης πληροφορίας και στεγανογραφίας, δεν υποκινεί υποψίες ότι πίσω από ένα συγκεκριμένο ορατό περιεχόμενο μπορεί να κρύβεται και ΜΗ ορατή πληροφορία. Η υδατοσήμανση αποκρύβει ένα μυστικό μήνυμα μέσα στο αρχικό περιεχόμενο για την προστασία του copyright ενός προϊόντος και την απόδειξη της αυθεντικότητάς ενώ παρέχει επίσης μηχανισμούς εξακρίβωσης του μυστικού περιεχομένου για την απόδειξη της υποκλοπής. Η βασική διαφορά μεταξύ στεγανογραφίας και υδατοσήμανσης αφορά το σκοπό της χρήσης τους. Στη στεγανογραφία, ο μη εξουσιοδοτημένος χρήστης προσπαθεί να αποκαλύψει αν μεταφέρεται κρυφή πληροφορία. Στην υδατοσήμανση, ο μη εξουσιοδοτημένος χρήστης προσπαθεί να αφαιρέσει το υδατογράφημα ώστε να τα αναπαράγει παράνομα στη συνέχεια χωρίς να μπορεί να ιχνηλατηθεί.

Οι διαδικασίες προσθήκης και ανίχνευσης υδατογραφήματος σε μια ΒΔ είναι αντίστροφες και αποτυπώνονται στην Εικόνα 4.



Εικόνα 4: Η διαδικασία της Δημιουργίας και της Ανίχνευσης ενός αόρατου υδατογραφήματος

Η δημιουργία ενός αόρατου υδατογραφήματος βασίζεται σε μια ομάδα παραμέτρων και ένα μυστικό κλειδί. Η υδατογραφημένη βάση μπορεί να διανεμηθεί, να αντιγραφεί νόμιμα ή παράνομα, όμως το κρυφό υδατογράφημα πάντα θα υπάρχει ενσωματωμένο στα δεδομένα. Οι επιθέσεις μπορούν να το αλλοιώσουν σε ένα βαθμό προσθέτοντας θόρυβο ή με αλλαγές στα δεδομένα της βάσης. Η ανίχνευση χρησιμοποιεί τις ίδιες παραμέτρους και το ίδιο κλειδί με στόχο να επιβεβαιώσει την ύπαρξη της υδατοσήμανσης με υψηλό βαθμό βεβαιότητας. Ο Πίνακας 5 συνοψίζει τους βασικούς τύπους υδατογραφημάτων με τις χρήσεις τους.

ΑΑ	***Τύπος***	***Χρησιμότητα***
1	*Ορατό*	*Δηλώνει φανερά ότι η εμπορική χρήση του αντικειμένου έχει νομικούς περιορισμούς και προστατεύεται από τη νομοθεσία της πνευματικής ιδιοκτησίας. Αποτρέπει την παράνομη αντιγραφή από ανυποψίαστους χρήστες.*
2	*Αόρατο*	*Ανίχνευση και εντοπισμός παράνομης εμπορικής συναλλαγής ή απόδειξη της πνευματικής ιδιοκτησίας. Έχει ως χαρακτηριστικό την ανθεκτικότητά σε επιθέσεις που αποσκοπούν στην εξάλειψη της υδατοσήμανσης.*
3	*Ανθεκτικό*	*Εντοπισμός παράνομης εμπορικής εκμετάλλευσης που προστατεύεται από το νόμο της πνευματικής ιδιοκτησίας. Τα υδατογραφήματα αυτά είναι ανθεκτικά σε επιθέσεις και αξιοποιούνται στον εντοπισμό του υδατογραφήματος ανεξαρτήτως της επεξεργασίας που έχει υποστεί το περιεχόμενο. Σε περίπτωση μη εντοπισμού του υδατογραφήματος, δηλαδή αν έχει επέλθει η αφαίρεση ή η διαστρέβλωσή του, τότε το περιεχόμενο ίσως να έχει αλλοιωθεί σε μεγάλο βαθμό. Υπάρχει λεπτή ισορροπία ανάμεσα στην ανθεκτικότητα και στην αδιορατότητα.*
4	*Μη ανθεκτικό (εύθραστο)*	*Σε περίπτωση υποψίας αλλοίωσης του περιεχομένου, η αλλοίωση αυτή αποδεικνύεται μέσω του υδατογραφήματος. Αν κάποιος κακόβουλος αλλοιώσει το περιεχόμενο, αλλοιώνει και το αόρατο υδατογράφημα, που αποδεικνύει ότι το περιεχόμενο δεν είναι πρωτότυπο αλλά έχει υποστεί επεξεργασία. Η μέθοδος αυτή ιχνηλατεί τον κακόβουλο και έχει χρησιμοποιηθεί σε νομικές υποθέσεις.*
5	*Αποτύπωμα*	*Ενσωματώνοντας αναγνωριστικά στο περιεχόμενο, επιτρέπεται η ανίχνευση του παραλήπτη. Ο πωλητής ενός αντικειμένου, για παράδειγμα, μπορεί να εισάγει μία μοναδική και αόρατη ετικέτα στο αντικείμενο, με στόχο την ταυτοποίηση με το πρόσωπο στο οποίο πουλήθηκε το αντικείμενο. Κάποια άλλη χρονική στιγμή ο πωλητής μπορεί να διαπιστώσει την ύπαρξη ενός παράνομου αντιγράφου του αντικειμένου αυτού. Το υδατογράφημα θα υποδείξει αυτόν στον οποίο πουλήθηκε το αντικείμενο και στην πιθανή δημιουργία του παράνομου αντιγράφου.*

Πίνακας 5: Οι βασικοί τύποι υδατογραφημάτων

Μια ταξινόμηση όλων των μεθόδων υδατοσήμανσης που έχουν εφαρμοσθεί σε βάσεις δεδομένων φαίνεται στην Εικόνα 5 με βάση την ταξινόμηση της πηγής [26].

Watermarking and Fingerprinting techniques	*BIT Resetting*	*Bit Pattern Based*
		Image Based
		Fingerprinting Based
	Data statistics modifying	*Bit Pattern Based*
		Image Based
	Constrained data content modifying	*Tuple Based*
		Attribute Based
		Image Based

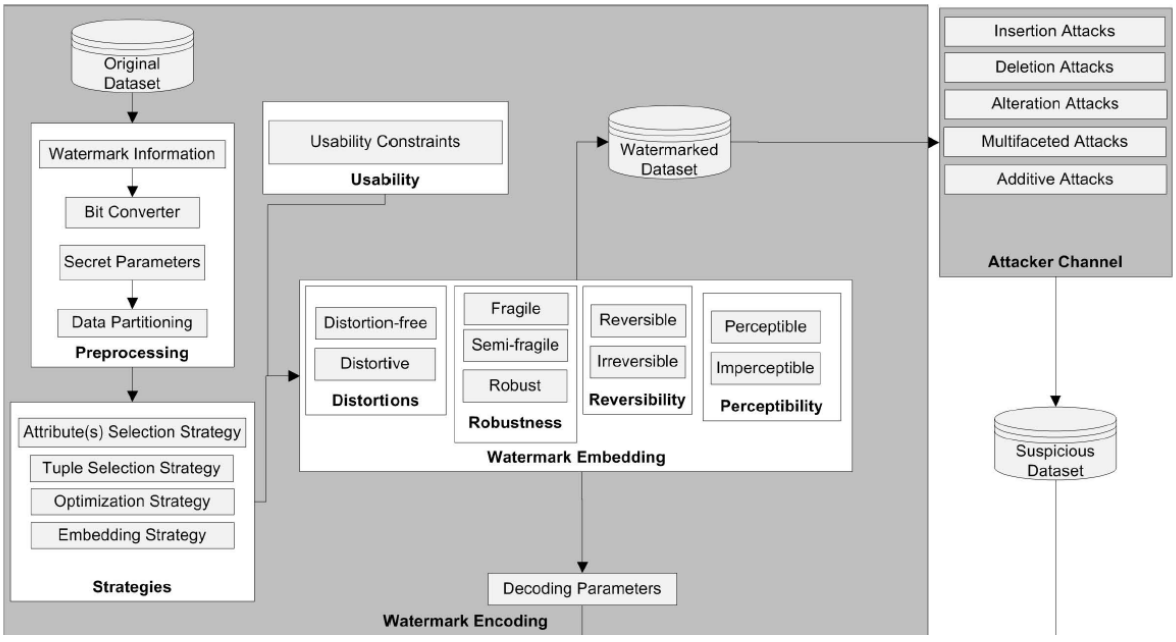
Εικόνα 5: Κατηγοριοποίηση των μεθόδων υδατογράφησης βάσεων δεδομένων (πηγή [26])

Στο ίδιο άρθρο [26] αποτυπώνονται διαγραμματικά η αρχιτεκτονική για την ενσωμάτωση και την εξαγωγή των υδατογραφημάτων. Οι τρείς δημοφιλείς τεχνικές που αναφέρονται είναι :

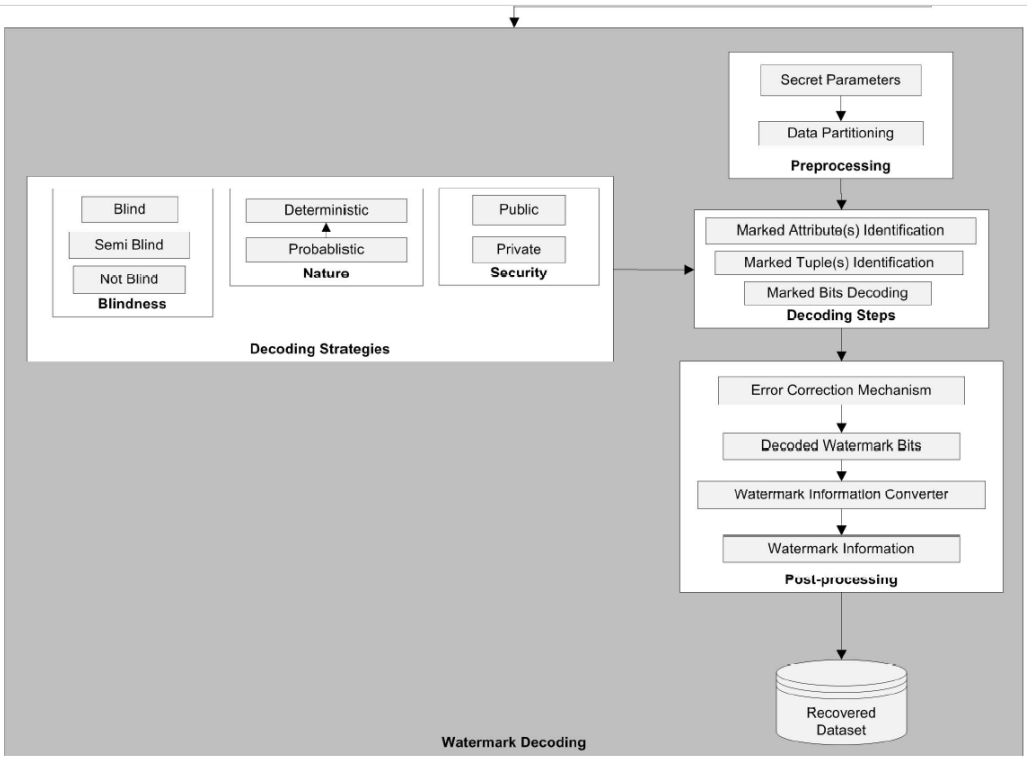
i. Η αντικατάσταση επιλεγμένων δυαδικών ψηφίων - Bit Resetting techniques (BRT)

ii. Η χρήση στατιστικών στοιχείων περιεχομένου (μέσος, διακύμανση, κατανομή) - Data statistics modifying techniques (DSMT)

iii. Σχετικές με μεταβολές που αξιοποιούν την σειρά των πλειάδων εισάγοντας ειδικούς χαρακτήρες σε συμβολοσειρές - Constrained data content modifying techniques (CDCMT)



Εικόνα 6: Η αρχιτεκτονική ενσωμάτωσης υδατοσήμανσης σε σύνολα δεδομένων (πηγή [26])



Εικόνα 7: Η αρχιτεκτονική ανίχνευσης υδατοσήμανσης και ανάκτησης αρχικών δεδομένων (πηγή [26])

Μια αναλυτικότερη ταξινόμηση, οδηγεί στις παρακάτω υποκατηγορίες μεθόδων υδατοσήμανσης [26].

Τεχνικές παρέμβασης σε δυαδικά ψηφία (bit resetting) :

i. Bit pattern based bit-resetting (BPBBRT): Χρήση ενός μοτίβου από δυαδικά ψηφία (bits) ως υδατόσημο που ενσωματώνεται στην πληροφορία.
ii. Image based bit-resetting (IBBRT): Χρήση μιας εικόνας ως υδατόσημο.
iii. Bit-resetting fingerprinting (BRFT): Χρήση μοτίβου από bits στη δημιουργία ψηφιακών αποτυπωμάτων.

Τεχνικές επαναφοράς δυαδικών ψηφίων (bit restoring) :

i. Τεχνικές επαναφοράς δυαδικών ψηφίων που βασίζονται σε μοτίβα bit (BPBBRT): ένα μοτίβο bit χρησιμοποιείται ως υδατογράφημα και ενσωματώνεται στα επιλεγμένα bit των δεδομένων.
ii. Τεχνικές επαναφοράς bit βάσει εικόνας (IBBRT): μια εικόνα χρησιμοποιείται ως υδατογράφημα, για ενσωμάτωση στα επιλεγμένα bit των δεδομένων.
iii. Τεχνικές δακτυλικών αποτυπωμάτων επαναφοράς bit (BRFT): χρησιμοποιούν δακτυλικά αποτυπώματα (αναγνωριστικά) για τη σήμανση επιλεγμένων bits των σχεσιακών βάσεων δεδομένων.

Στατιστικές τεχνικές παρέμβασης (data statistics and modifying)

i. Bit pattern based data statistics-modifying (BPBDSMT): Χρήση μοτίβων bits για ενσωμάτωση στα στατιστικά των δεδομένων.
ii. Image based data statistics-modifying techniques (IBDSMT): Χρήση εικόνας στην ενσωμάτωση υδατοσήμανσης σε στατιστικά δεδομένων.

Constrained Data Content-Modifying Techniques

i. Tuple based constrained data content-modifying (TBCDCMT): μεταβολή του περιεχομένου σε επίπεδο πλειάδας.
ii. Attribute based constrained data content-modifying (ABCDCMT): μεταβολή του περιεχομένου σε επίπεδο αναγνωριστικού.
iii. An image based constrained data content-modifying (IBCDCMT): χρήση εικόνας για τη υδατοσήμανση του περιεχομένου.

Ο Πίνακας 6 έχει αντληθεί χωρίς επεξεργασία από την πηγή [26], επειδή συνοψίζει όλες τις παραπάνω τεχνικές. Οι στήλες με επικεφαλίδες R1 έως R5 αναφέρονται στην ανθεκτικότητα (robustness) των τεχνικών υδατογράφησης και αντιστοιχούν σε επιθέσεις : εισαγωγής (R1), διαγραφής (R2), αλλαγής (R3), μίγματος (R4) και ψεύτικου υδατογραφήματος (R5).

Class of Techniques	R_1	R_2	R_3	R_4	R_5
BRT	Usually resilient except techniques that use specific order (or position) of tuples for watermark insertion.	Usually resilient except techniques that: use specific order (or position) of tuples for watermark insertion.	Usually not resilient because small manipulations can easily disturb the marked bits and even multi-bit watermarks can be corrupted.	Usually not resilient due to facts depicted in R_1, R_2, and R_3.	Usually not resilient apart from techniques that are specifically designed for tackling attacks of type A_5 and they still have limitations.
DSMT	Usually resilient except techniques that use specific order (or position) of tuples for watermark insertion.	Usually resilient except techniques that use specific order (or position) of tuples for watermark insertion.	Usually resilient.	Usually techniques that mark more tuples are more robust to such attacks. But still this type of attack needs to be studied in depth because most of the techniques proposed so far do not discuss this attack.	Usually not resilient apart from techniques that are specifically designed for tackling attacks of type A_5 and they still have limitations.
CDCMT	Usually resilient except techniques that use specific order (or position) of tuples or tuples for watermark insertion.	Usually not resilient because deletion attack may cause deletion of newly inserted content if new content is inserted during watermark embedding.	Usually not resilient because small manipulations can easily disturb the marked bits without degrading the data usability significantly.	Usually not resilient due to the facts depicted in R_1, R_2, and R_3.	Usually not resilient apart from techniques that are specifically designed for tackling attacks of type A_5 and they still have limitations.

Πίνακας 6: Η ανθεκτικότητα γνωστών κατηγοριών υδατογράφησης (πηγή [26])

Από την παραπάνω σύγκριση, φαίνεται ότι οι τεχνικές DSMT, υπερτερούν αλλά έχουν περιορισμούς όταν υπάρχει προϋπόθεση συγκεκριμένης ταξινόμηση.

Μια διαφορετική ταξινόμηση φαίνεται παρακάτω.

Πληροφορία Υδατοσήμανσης	**Τύπος Δεδομένων**	**Επίπεδο αλλοίωσης**	**Ορατότητα**	**Σκοπός**
Μοτίβο δυαδικών στοιχείων	Αριθμητικά	Πολλαπλά bits	Αόρατο	Απόδειξη ιδιοκτησίας
Εικόνα	Αλφαριθμητικά	Bit ή χαρακτήρας ή χαρακτηριστικό	Αόρατο	Απόδειξη ιδιοκτησίας ή ακεραιότητα
Ηχητικό μήνυμα	Αριθμητικά	Πολλαπλά bits	Αόρατο	Απόδειξη ιδιοκτησίας
Περιεχόμενο ΒΔ	Αριθμητικά	Πολλαπλά bits	Αόρατο	Απόδειξη ιδιοκτησίας
Συμβολοσειρά	Κατηγοριοποιημένα	Bit	Αόρατο	Απόδειξη ιδιοκτησίας
Αποτύπωμα	Αριθμητικά	Πολλαπλά bits	Αόρατο	Ιχνηλάτιση πειρατή

Πίνακας 7 : Είδη υδατογράφησης και αποτυπωμάτων (πηγή [26])

Συνδυάζοντας την ταξινόμηση των μεθόδων που αναφέρεται στην πηγή [41], ο Πίνακας 8 συνοψίζει τα μέσα, τις ιδιότητες και τους σκοπούς που εξυπηρετεί ένα υδατογράφημα.

Μέσα	**Ιδιότητες**	**Σκοποί**
Κείμενο	Τυφλό ή ΟΧΙ	Αναγνώριση δημιουργού
Ήχος	Δημόσιο ή Ιδιωτικό	Αποτροπή παράνομης αντιγραφής
Εικόνα	Ορατό ή Αόρατο	Αυθεντικοποίηση
Κινούμενη εικόνα	Ισχυρό ή Εύθραυστο	Ακεραιότητα
Γραφικά	Αναστρέψιμο ή ΜΗ	Αποτύπωμα
Βάσεις Δεδομένων	Προσαρμόσιμο στα δεδομένα	Έλεγχος της αντιγραφής

Πίνακας 8: Ταξινόμηση των μεθόδων υδατογράφησης με βάση το μέσο, τις ιδιότητες και το σκοπό

3.4 Τα ψηφιακά συστήματα διαχείρισης δικαιωμάτων

Τα Συστήματα Διαχείρισης Πνευματικών Δικαιωμάτων (DRM) υποστηρίζουν τη διανομή περιεχομένου και περιλαμβάνουν συστήματα αναγνώρισης και ανταλλαγής πληροφοριών πνευματικής ιδιοκτησίας, ενώ καλύπτουν απαιτήσεις διαλειτουργικότητας, ασφάλειας, ευκολίας στη χρήση και υπηρεσίες εξουσιοδότησης. Η διαχείριση ψηφιακών δικαιωμάτων (DRM) αφορά τον έλεγχο και τη διαχείριση της πρόσβασης σε πνευματικό έργο που προστατεύεται από πνευματικά δικαιώματα με τη χρήση τεχνολογικών μέσων. Το DRM δίνει τη δυνατότητα στους δημιουργούς να παρέχουν εξουσιοδότηση σε συσκευές και χρήστες ώστε να προστατεύουν τα έργα που υπόκεινται σε προστασία, καθιστώντας αδύνατη την κλοπή ή την παράνομη χρήση των έργων αυτών [79]. Μπορεί να περιλαμβάνεται η χρήση κωδικών που απαγορεύει την αντιγραφή περιεχομένου ή περιορίζει τον αριθμό των συσκευών από τις οποίες μπορεί να προσπελαστεί ένα προϊόν. Ο Πίνακας 9 αναφέρει κάποια παραδείγματα γνωστών DRMs.

ΑΑ	**Περίπτωση**	**Περιγραφή**
1	Κατάστημα Apple iTunes	Περιορίζει τον αριθμό των συσκευών που μπορούν να χρησιμοποιούν οι πελάτες για να ακούν τραγούδια. Τα αρχεία ήχου του iTunes έχουν πληροφορία σχετική με την αγορά και τη χρήση των τραγουδιών
2	Κατάστημα Apple iBooks	Προστατεύει το περιεχόμενο με τεχνολογία FairPlay που διασφαλίζει ότι τα βιβλία μπορούν να διαβαστούν μόνο σε συσκευές iOS.
3	Spotify	Αξιοποιεί τεχνολογία blockchain για να επιτρέψει την πληρωμή των μουσικών καλλιτεχνών μέσω κρυπτονομισμάτων
4	Microsoft	Όποιος κατεβάζει λογισμικό της Microsoft, όπως προγράμματα Windows ή Office, πρέπει να αποδεχτεί την άδεια χρήσης και να εισάγει ένα κλειδί ή να το ανακτήσει μέσω διαδικτύου πριν για την εγκατάσταση.
5	Ευαίσθητα έγγραφα	Πολλοί οργανισμοί προστατεύουν εμπιστευτικά δεδομένα, επιχειρηματικά σχέδια και συμβάσεις και παρακολουθούν την πρόσβαση στα αρχεία τους.
6	Συμμόρφωση με κανονισμούς	Επιτρέπει σε οργανισμούς υγειονομικής περίθαλψης να πληρούν τις απαιτήσεις του νόμου περί φορητότητας και λογοδοσίας ασφάλισης υγείας (HIPAA), τους βοηθά στην κάλυψη των αναγκών του Νόμου περί

ΑΑ	Περίπτωση	Περιγραφή
		Απορρήτου των Καταναλωτών της Καλιφόρνια (CCPA) και του Γενικού Κανονισμού Προστασίας Δεδομένων (GDPR) της ΕΕ.

Πίνακας 9: Παραδείγματα γνωστών DRM

Οι πλατφόρμες DRM επιτρέπουν στους οργανισμούς να ελέγχουν το υλικό τους που προστατεύεται από πνευματικά δικαιώματα με μια σειρά από τεχνολογίες :

i. Υδατογράφημα: Επιτρέπει στους δημιουργούς ή στους κατόχους πνευματικών δικαιωμάτων να παρακολουθούν τη μη εξουσιοδοτημένη χρήση.

ii. Μεταδεδομένα: Επιτρέπουν την καταγραφή και παρακολούθηση των πληροφοριών πνευματικών δικαιωμάτων και αδειών χρήσης τους.

iii. Ενσωμάτωση κωδικών: Οι κάτοχοι περιεχομένου μπορούν να χρησιμοποιούν κωδικούς ενσωμάτωσης που ελέγχουν τις δημοσιεύεις των έργων τους στο διαδίκτυο. Αυτή η διαδικασία ονομάζεται "δημιουργία μία φορά, δημοσίευση παντού" (COPE).

iv. Συμφωνίες άδειας χρήσης: Προστατεύουν περιεχόμενο και λογισμικό απαιτώντας από τους χρήστες να συμφωνούν στην άδειας χρήσης (EULA) την πρώτη φορά πρόσβασης στο σχετικό ιστότοπο.

Οι λειτουργίες που παρέχει ένα DRM περιλαμβάνουν [78, 80] την αποτροπή πειρατείας σε πολύτιμα έργα πνευματικής ιδιοκτησίας, τη διαχείριση της εξουσιοδοτημένης πρόσβασης , την εφαρμογή περιορισμών περιεχομένου (εκτύπωση, αντιγραφή, υδατοσήματα, δεδομένα, συσκευές) και την παρακολούθηση ενεργειών των εγγεγραμμένων χρηστών μέσα από λεπτομερείς αναφορές. Οι προδιαγραφές ενός συστήματος DRM αναλύονται στην πηγή [77] και έχουν οργανωθεί στον παρακάτω πίνακα.

ΑΑ	Απαίτηση	Περιγραφή
1	Μόνιμη προστασία	Οι έλεγχοι πρόσβασης που επιβάλλονται από τον κάτοχο των δικαιωμάτων πρέπει να επιβάλλονται ανεξάρτητα από τη θέση δημοσίευσης των ψηφιακών δεδομένων. Σε επίπεδο συσκευής υλοποιείται με το να μην μπορεί να διαβάζεται το αρχείο.
2	Διεταιρικές Συναλλαγές	Αφορά τη σύναψη συνεργασιών για τη δημιουργία νέων προϊόντων και υπηρεσιών που περιλαμβάνουν ρήτρες εμπιστευτικότητας και ασφαλή επικοινωνία.
3	Φορητότητα	Έχει διαφορετικές έννοιες και περιλαμβάνει υποκατηγορίες : - Χρονική μετατόπιση που αφορά στον έλεγχο της πρόσβασης προστατευμένων δεδομένων σε συγκεκριμένες ώρες. - Χωρική μετατόπιση ή μετατόπιση πλατφόρμας που αναφέρεται στην πρόσβαση από διαφορετικά λειτουργικά συστήματα ή συσκευές. - Μετατόπιση μορφής αφορά την αλλαγή της μορφής των αρχείων όταν αυτά διατίθενται στο κοινό.
4	Εξαγωγή ή Απόσπαση	Επιτρέπει σε έναν χρήστη να λάβει ένα συγκεκριμένο τμήμα δεδομένων από μια πηγή για να το ενσωματώσει σε άλλο αρχείο

ΑΑ	Απαίτηση	Περιγραφή
5	Ενοποίηση με εφαρμογές	Αφορά λογισμικό που έχει αναπτυχθεί για να καλύψει ειδικές ανάγκες επεξεργασίας εμπιστευτικών εγγράφων.
6	Μεταβίβαση Δικαιωμάτων	Καλύπτουν τις περιπτώσεις μετακίνησης εργαζομένων, αλλαγων domain names, καταχωρημένων διευθύνσεων στο διαδίκτυο και δικαιωμάτων πρόσβασης. Λόγω της ευκολίας αναπαραγωγής των ψηφιακών δεδομένων, τα προστατευόμενα δεδομένα δεν θα πρέπει να είναι προσβάσιμα μετά τη μεταβίβαση του δικαιώματος.
7	Διαχείριση δικαιωμάτων πρόσβασης	Τα αρχικά δικαιώματα μπορεί να είναι πολύ περιοριστικά ή υπερβολικά και να χρειάζονται ρύθμιση μετά τη διανομή των προστατευόμενων δεδομένων.
8	Παρακολούθηση	Αφορά την παρακολούθηση της πρόσβασης σε εμπιστευτικά δεδομένα μέσω αρχείων καταγραφής.
9	Χρήση εκτός σύνδεσης	Καλύπτουν περιπτώσεις μη πρόσβασης στο διαδίκτυο και χρήσης αντιγράφων (offline) από εργαζόμενους. Αυτό μειώνει τις δυνατότητες παρακολούθησης και περιορίζει τον έλεγχο.
10	Αυθεντικοποίηση Εξουσιοδότηση	Δυνατότητα αναγνώρισης και συσχέτισης χρηστών με τα δικαιώματα πρόσβασης στα έργα και μηχανισμοί συσχέτισης των κατόχων δικαιωμάτων.
11	Επαλήθευση	Αφορά το να μπορούν οι έντιμοι χρήστες να αποδείξουν εύκολα ότι έχουν νόμιμη πρόσβαση στο προστατευόμενο έργο.
12	Ασφάλεια	Περιλαμβάνει όλες τις απαιτήσεις της ασφάλειας : Ακεραιότητα, Εμπιστευτικότητα, Διαθεσιμότητα, Αυθεντικοποίηση, ΜΗ αποποίηση ευθύνης

Πίνακας 10 : Προδιαγραφές συστημάτων DRM

Οι τεχνικές ασφαλείας που εφαρμόζονται από τα DRM [78] περιλαμβάνουν :

i. Κρυπτογραφία στην προστασία του περιεχομένου από την πρόσβαση και την ασφάλεια των επικοινωνιών μεταξύ του χρήστη και του διανομέα.

ii. Αναγνωριστικά περιεχομένου (ID) για την ιχνηλάτιση και την ανίχνευση της πειρατείας. Το σχήμα Digital Object Identification (DOI) λειτουργεί παρόμοια με έναν γραμμωτό κώδικα, αξιοποιεί στατικές διευθύνεις URL, αναζητώντας την τρέχουσα θέση του περιεχομένου για να ανακατευθύνει σε αυτήν.

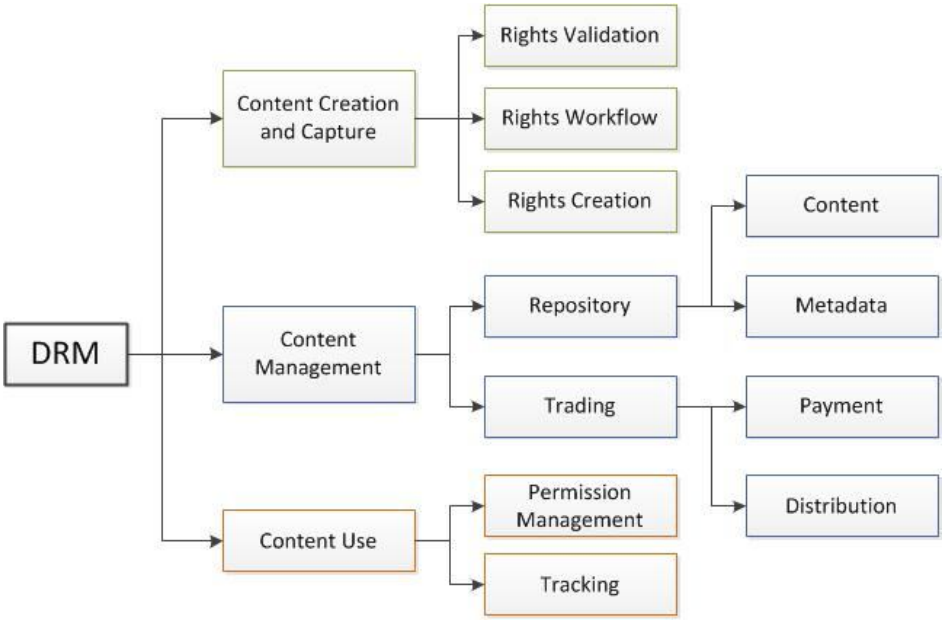
iii. Ψηφιακό ορατό αποτύπωμα μπορεί να είναι ένα μικρό δείγμα του περιεχομένου ή ένα αόρατο υδατογράφημα που ενσωματώνεται στο ψηφιακό περιεχόμενο.

iv. Ιχνηλασία μέσω υδατογράφησης για τον προσδιορισμό του νόμιμου χρήστη μέσω του αντιγράφου που διέρρευσε.

v. Ενημέρωση και διαλειτουργικότητα που διασφαλίζουν την ανθεκτικότητα των συστημάτων DRM και τη συνεργασία τους με αντίστοιχα συστήματα μέσα από πρωτόκολλα, τυποποίηση γλωσσών και αδειών.

vi. Χρήση μεταδεδομένων και προτύπων, κεφαλίδες αρχείων πολυμέσων, ενσωμάτωση λειτουργιών σε λειτουργικά συστήματα, συστήματα μετάδοσης και διανομής online περιεχομένου live streaming.

Η λειτουργική αρχιτεκτονική ενός DRM [81,82] για τη δημιουργία συστημάτων με δυνατότητα ψηφιακών δικαιωμάτων φαίνεται στην
Εικόνα 8.



Εικόνα 8: Η λειτουργική αρχιτεκτονική ενός DRM (πηγή [81])

Οι βασικές λειτουργίες περιλαμβάνουν τη φύλαξη, τη δημιουργία και την αποτύπωση των περιουσιακών στοιχείων πνευματικής ιδιοκτησίας, τη διαχείριση της δημιουργία περιεχομένου, τη διεκδίκηση δικαιωμάτων όταν το περιεχόμενο δημιουργείται για πρώτη φορά, τη χρήση των περιουσιακών στοιχείων (για εμπορία ή πληρωμή), τους περιορισμούς υποστήριξης σε συγκεκριμένα συστήματα, τη διαχείριση δικαιωμάτων, τη διανομή και την ιχνηλάτιση.

4 Προσεγγίσεις υδατογράφησης και επιθέσεων

4.1 Οι προσεγγίσεις υδατογράφησης

Οι τρείς γνωστές προσεγγίσεις υδατοσήμανσης αριθμητικών δεδομένων είναι :

i. Η ισχυρή υδατοσήμανση που στοχεύει στην αναγνώριση του δημιουργού και στην προστασία των πνευματικών δικαιωμάτων του. Είναι αρκετά ανθεκτική σε επιθέσεις αφαίρεσης ή καταστροφή της υδατοσήμανσης

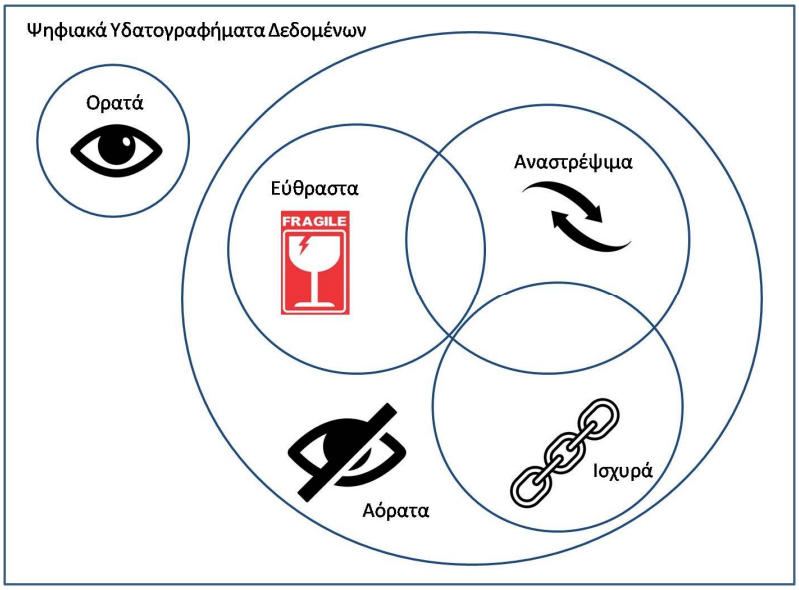
ii. Η εύθραυστη υδατοσήμανση που στοχεύει στην επαλήθευση της ακεραιότητας των δεδομένων αφού η σήμανση αλλοιώνεται σε οιαδήποτε προσπάθεια μη εξουσιοδοτημένης τροποποίησης των δεδομένων.

iii. Το ψηφιακό αποτύπωμα που αποσκοπεί στην ιχνηλάτιση και ενσωματώνεται σε κάθε διανεμημένο αντίγραφο.

ΑΑ	*Κατηγορία*	*Στόχος υδατοσήμανσης*	*Στόχος επίθεσης*
1	*Ισχυρό*	*Αναγνώριση δημιουργού*	*Αλλοίωση ή αφαίρεση του watermark*
2	*Εύθραυστο*	*Ακεραιότητα*	*Τροποποίηση δεδομένων χωρίς αλλοίωση του watermark*
3	*Αποτύπωμα*	*Ιχνηλάτηση πειρατή*	*Προσθήκη νέου ισχυρού watermark και πιθανή αλλοίωση του αρχικού watermark*

Πίνακας 11: Κατηγοριοποίηση προσεγγίσεων υδατοσήμανσης δεδομένων

Υπάρχουν διαδικασίες υδατογράφησης που επιτρέπουν την επαναφορά των δεδομένων στην αρχική κατάσταση (αντιστρέψιμες) [27,36,86]. Με βάση την παραπάνω διάκριση, η Εικόνα 9 διακρίνει τις διαφορετικές προσεγγίσεις.



Εικόνα 9: Ορατά, αόρατα, ισχυρά, εύθραστα και αντιστρέψιμα υδατογραφήματα

Κάθε προσέγγιση περιλαμβάνει μια σειρά από ιδιότητες, παραμέτρους, αλγόριθμους και πειραματικές υλοποιήσεις της βιβλιογραφίας. Οι ιδιότητες που διακρίνουν ένα ισχυρό υδατογράφημα [21, 76] είναι :

i. Ανθεκτικότητα σε επιθέσεις μεταβολών τιμών
ii. Ακρίβεια στην ανίχνευση
iii. Ανανέωση μετά από ενέργειες τροποποιήσεων των δεδομένων
iv. Τυφλή επιβεβαίωση με ανίχνευση χωρίς πρόσβαση στον αρχικό σύνολο
v. Δημόσιοι αλγόριθμοι με χρήση μυστικών κλειδιών
vi. Χρηστικότητα των δεδομένων μετά από ελεγχόμενη αλλοίωση

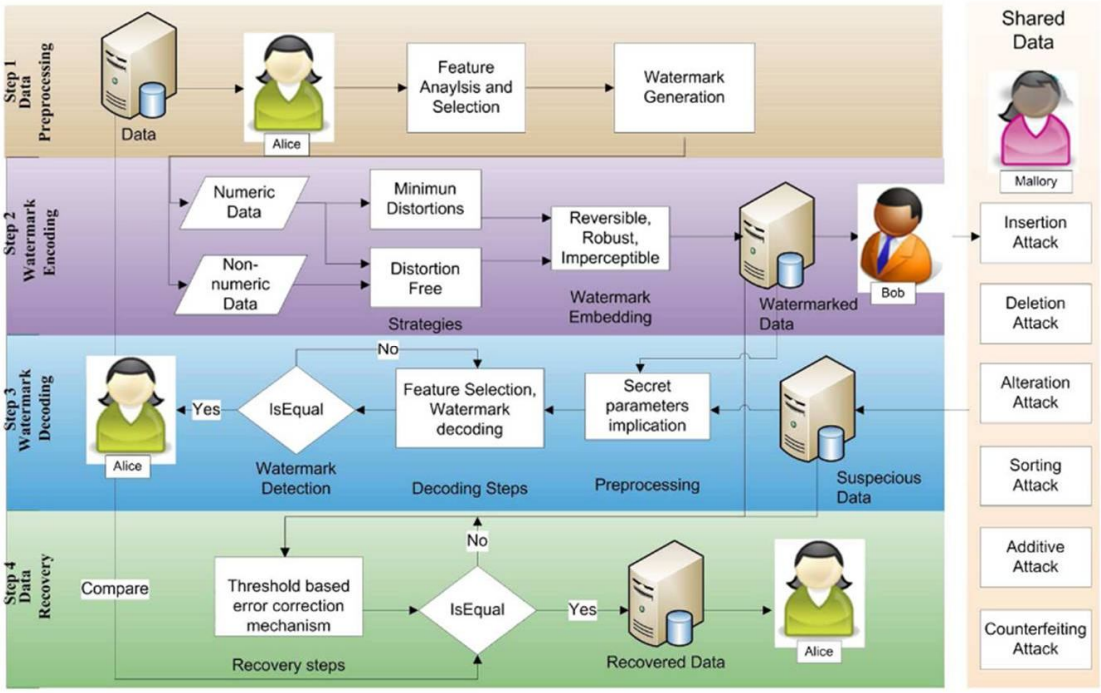
Η εργασία [76] περιλαμβάνει μια συγκριτική ποσοτική αξιολόγηση επτά διαφορετικών προσεγγίσεων ισχυρής υδατοσήμανσης, με βάση τα παραπάνω έξι κριτήρια. Οι ιδιότητες που διακρίνουν ένα εύθραυστο υδατογράφημα [58] είναι :

i. Αορατότητα : καμία αλλοίωση των δεδομένων
ii. Ασφάλεια : προσθήκη και ανίχνευση με τη χρήση ιδιωτικού κλειδιού
iii. Τυφλή επιβεβαίωση : ανίχνευση χωρίς πρόσβαση στον αρχικό σύνολο
iv. Ανίχνευση των τροποποιήσεων του εισβολέα : κατά προσέγγιση προσδιορισμός

Οι ιδιότητες που διακρίνουν ένα ψηφιακό αποτύπωμα είναι οι ίδιες με αυτές του ισχυρού υδατογραφήματος αφού οι διαφορές μεταξύ των δυο αφορούν το σκοπό και όχι τα τεχνικά χαρακτηριστικά.

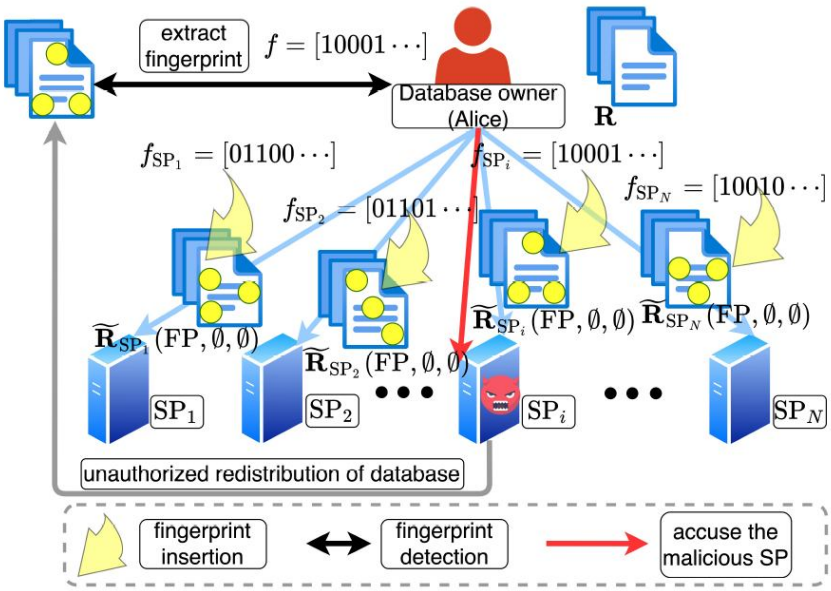
Η πηγή [86] αποτυπώνει τον κύκλο ζωής ενός αναστρέψιμου υδατογραφημένου συνόλου δεδομένων, που περιλαμβάνει τα τέσσερα στάδια της προεπεξεργασίας, της κωδικοποίησης, της αποκωδικοποίησης και της επαναφοράς. Περιλαμβάνει τη διάκριση μεταξύ αριθμητικών και ΜΗ δεδομένων καθώς και τη διάκριση μεταξύ σχημάτων που προκαλούν παραμόρφωση ή αφήνουν τα δεδομένα ανέπαφα. Οι ρόλοι που εμπλέκονται είναι ο νόμιμος ιδιοκτήτης ή κατασκευαστής

(Alice), ο νόμιμος αποδέκτης (Bob) και ο επιτιθέμενος/πειρατής που προσπαθεί να αλλοιώσει τα δεδομένα ώστε να αποτύχει η αποκρυπτογράφηση και η απόδειξη του πραγματικού ιδιοκτήτη.



Εικόνα 10: Ο κύκλος ζωής ενός υδατογραφημένου συνόλου δεδομένων (πηγή [86])

Ένα ψηφιακό αποτύπωμα έχει τα ίδια τεχνικά χαρακτηριστικά με ένα ισχυρό υδατογράφημα, μόνο που κάθε αντίγραφο υπογράφεται ψηφιακά για κάθε νόμιμο αποδέκτη με στόχο την ιχνηλάτισή του. Συστήματα όπως το Vanilla System [67, 69] υιοθετούν αυτή τη διαδικασία.



Εικόνα 11: Η χρήση ψηφιακού αποτυπώματος στην ελεγχόμενη διάθεση αντιγράφων (πηγή [67])

Στην

Εικόνα 11, ο ιδιοκτήτης ή κατασκευαστής μιας συλλογής δεδομένων (Alice) διαθέτει αντίγραφα των δεδομένων με ελεγχόμενο τρόπο σε παρόχους (Service Providers), αφού πρώτα προσθέσει ένα αποτύπωμα ή υδατογράφημα με διαφορετικό κλειδί για τον καθένα. Αν κάποιος πάροχος δημοσιεύει παράνομα (χωρίς εξουσιοδότηση) κάποια δεδομένα και αυτά τα αντίγραφα καταλήξουν στα χέρια του ιδιοκτήτη και ελεγχθούν με εξαγωγή του αποτυπώματος, ο κακόβουλος μπορεί να κατηγορηθεί άμεσα για την ΜΗ εξουσιοδοτημένη διανομή.

4.2 Οι επιθέσεις στα υδατογραφήματα

Ένα υδατογράφημα σε αριθμητικά δεδομένα, συνήθως προσθέτει θόρυβο αλλάζοντας ένα αριθμό δυαδικών ψηφίων (των λιγότερο σημαντικών) με ελεγχόμενο τρόπο ώστε να μπορεί να ανιχνευθεί στο μέλλον. Έτσι η παρέμβαση είναι αόρατη, μυστική και επηρεάζει τη χρησιμότητα των δεδομένων σε ένα μικρό βαθμό.

Most Significant Bits (MBS)								**…**	**Least Significant Bits (LSB)**			
0	1	0	1	1	0	0	1		1	0	1	1

Οι επιθέσεις έχουν ως στόχο την εξάλειψη του υδατογραφήματος, οπότε επικεντρώνονται με διάφορους τρόπους στην αλλοίωση των LSBs που περιέχουν το κρυμμένο υδατογράφημα. Οι πιο γνωστές επιθέσεις που αναφέρονται στη βιβλιογραφία [28, 36, 41, 47, 52, 74] είναι :

Επίθεση δυαδικού ψηφίου με παραλλαγές Bit0, Bit1, BitFlip, BitRandom όπου τυχαία bits αριθμητικών δεδομένων μετατρέπονται σε 0 σε 1 ή αντιστρέφονται οι τιμές τους. Η επίθεση bit-flipping αναστρέφει τα LSB των τυχαία επιλεγμένων αριθμητικών ιδιοτήτων και πλειάδων.

Επίθεση στρογγυλοποίησης που επιλέγει το βαθμό στρογγυλοποίησης σε επίπεδο δεκαδικού ψηφίου αλλοιώνοντας το υδατογράφημα και τα δεδομένα. Ο βαθμός στρογγυλοποίησης είναι κρίσιμος εδώ, ένας μεγάλος βαθμός αλλοιώνει τα δεδομένα άσκοπα, ένας μικρός βαθμός δεν αλλοιώνει το υδατογράφημα αρκετά.

Επίθεση γραμμικού μετασχηματισμού που μετατρέπει αριθμητικά δεδομένα σε άλλη κλίμακα (πχ Celsius σε Fahrenheit, χιλιόμετρα σε μίλια). Η επίθεση γραμμικού μετασχηματισμού είναι στοχευμένη και περιλαμβάνει τη μετατροπή μονάδων μέτρησης. Στη φυσική έχει υιοθετηθεί το SI από το 1960 (International System of Units) για τις βασικές μονάδες μέτρησης των κοινών φυσικών μεγεθών όπως ο χρόνος (second), το μήκος (meter), η μάζα (Kg) με αντίστοιχα (υπο)πολλαπλάσια καθώς και η θερμοκρασία (kelvin). Υπάρχουν ακόμα και σήμερα διαθέσιμοι πίνακες δεδομένων σε τοπικές ή παραδοσιακές μονάδες μέτρησης μήκους (inch, foot, yard, mile), όγκου (gallon, quart, pint), μάζας (pound, ounce) και θερμοκρασίας (Fahrenheit, Celcius) σε χώρες όπως η Αμερική και η Αγγλία. Οι σχέσεις ανάμεσα στις μονάδες μέτρησης U_1,U_2 είναι γραμμικές (της μορφής $U_1=a*U_2+b$) ενώ υπάρχουν εργαλεία και πίνακες μετατροπής τους στο διαδίκτυο [www.unitconverters.net]. Η μετατροπή των μεγεθών δεν αλλοιώνει τα δεδομένα και δεν κινεί υποψίες, όμως θα πρέπει η επικεφαλίδα να αναφέρει την ακριβή μονάδα της μέτρησης.

Επίθεση διαγραφής ή προσθήκης πλειάδων σε τυχαίες θέσεις, επιφέροντας μεγάλη αλλοίωση στα δεδομένα ειδικά αν αφορά πολλές πλειάδες. Το υδατογράφημα δεν μπορεί να ανακτηθεί εάν μια πλειάδα που περιέχει πληροφορίες υδατογραφήματος διαγραφεί. Έτσι, η

επίθεση διαγραφής έχει τη χειρότερη επίδραση στον αριθμό των πλειάδων που έχουν υδατογραφηθεί.

Επίθεση αντιμετάθεσης ζευγών πλειάδων, που επηρεάζουν το υδατογράφημα το οποίο δημιουργήθηκε σε απόλυτες θέσεις με δεδομένη ταξινόμηση.

Επίθεση ταξινόμησης πλειάδων, που επαναταξινομούν τις πλειάδες με βάση ένα ή περισσότερα χαρακτηριστικά, καταστρέφοντας το υδατογράφημα της αρχικής ταξινόμησης

Επίθεση διεκδίκηση της ιδιοκτησίας με την εισαγωγή πλαστού υδατοσήματος. Αν το νόμιμο υδατόσημα παραμείνει παράλληλα με το παράνομο, είναι δύσκολη η διάκριση του νόμιμου ιδιοκτήτη. Υπάρχουν διάφοροι τρόποι αντιμετώπισης της υδατοσήμανσης σε ψευδείς ισχυρισμούς ιδιοκτησίας. Μια μέθοδος είναι να ζητηθεί και από τους δύο (τον ιδιοκτήτη και τον πειρατή) να παράγουν το πρωτότυπο αντίγραφο των δεδομένων πριν από την εισαγωγή του υδατογραφήματος. Ο νόμιμος κάτοχος μπορεί να αποδείξει την παρουσία του υδατογραφήματος στο πρωτότυπο ενώ ο πειρατής δεν μπορεί να δείξει την παρουσία του υδατογραφήματος στην αρχική βάση δεδομένων. Εναλλακτικά οι αξιώσεις ιδιοκτησίας μπορούν να επιλυθούν μόνο μέσα από κεντρικό μητρώο που το διαχειρίζεται μια αξιόπιστη τρίτη οντότητα. Ο νόμιμος κάτοχος Ν δημοσιεύει τα δεδομένα και προσαρτά το κλειδί στο μητρώο. Αν ο πειρατής Π διεκδικεί αργότερα την ιδιοκτησία της βάσης, η τρίτη οντότητα επιλύει τη διαφορά καθορίζοντας ότι το υδατογράφημα του Ν υπάρχει στα δεδομένα και το κλειδί του Ν προσαρτήθηκε στο μητρώο πριν από το κλειδί του Π. Εάν οι αξιώσεις ιδιοκτησίας δεν μπορούν να επιλυθούν, τότε οι πελάτες είναι επιφυλακτικοί στη χρήση και των 2 εκδόσεων των βάσεων.

Η ταξινόμηση, η αντιμετάθεση και ο μετασχηματισμός αλλοιώνουν μόνο το υδατογράφημα και όχι τα δεδομένα, η διεκδίκηση της ιδιοκτησίας, οι επιθέσεις bit και η στρογγυλοποίηση αλλοιώνουν δεδομένα και υδατογράφημα, ενώ η διαγραφή ή προσθήκη πλειάδων αλλοιώνει αρχικά τα δεδομένα και κατ'επέκταση και το υδατογράφημα σε μικρότερο βαθμό.

Η αγγλική ορολογία όλων των επιθέσεων που αναφέρονται στη βιβλιογραφία [21, 28, 36, 41, 47] και στοχεύουν στην αλλοίωση του υδατογραφήματος, συνοψίζεται παρακάτω.

ΑΑ	**Επίθεση**	**Περιγραφή και Χαρακτηριστικά**
1	Bit Attack	Αλλάζει bits των δεδομένων αλλοιώνοντας παράλληλα και τα δεδομένα.
2	Update	Προσθέτει, διαγράφει και εισάγει πλειάδες.
3	Zero Out Attack	Οι τιμές τυχαία επιλεγμένων bits μηδενίζονται.
4	One Out Attack	Οι τιμές τυχαία επιλεγμένων bits γίνονται 1.
5	Bit Flipping Attack	Οι τιμές τυχαία επιλεγμένων LSBs αντιστρέφονται σε τυχαίες πλειάδες.
6	Rounding Attack	Στρογγυλοποίηση των τιμών ενός χαρακτηριστικού.
7	Linear Transformation	Οι τιμές μετασχηματίζονται αλλάζει τιμές σε άλλη μονάδα μέτρησης.
8	Subset Attack	Ενημερώνει εγγραφές σε υποσύνολο πλειάδων ή χαρακτηριστικών.
9	Superset Attack	Προσθέτει νέες πλειάδες ή χαρακτηριστικά.
10	Collusion Attack	Αξιοποιεί πολλαπλά αντίγραφα με αποτυπώματα διαφορετικών αποδεκτών.
11	Mix and Match Attack	Ένωση πλειάδων από πολλαπλές σχέσεις που περιλαμβάνουν όμοιες πληροφορίες.
12	Majority Attack	Δημιουργεί αντίγραφο πίνακα όπου κάθε υπολογιζόμενο bit προκύπτει από την

ΑΑ	**Επίθεση**	**Περιγραφή και Χαρακτηριστικά**
		πλειοψήφησα τιμή του αντίστοιχου bit σε όλα τα διαθέσιμα αντίγραφα.
13	Claim of Ownership	Δημιουργεί 2^o ψεύτικο ανιχνεύσιμο υδατογράφημα και εισάγει αμφιβολίες ως προς τη νόμιμη ιδιοκτησία.
14	Additive Attack	Προσθήκη ψεύτικου υδατοσήματος σε νόμιμα υδατογραφημένο σύνολο για την διεκδίκηση της ιδιοκτησίας.
15	Invertibility Attack	Ανακάλυψη φανταστικού υδατογραφήματος από τυχαίο δείγμα υδατογραφημένης βάσης, ώστε στη φάση της ανίχνευσης να βγάλει αποτέλεσμα ψευδώς θετικό.
16	Subset Reverse Order	Ανταλλαγή της ταξινόμησης η της σειράς των πλειάδων ή των χαρακτηριστικών
17	Brute Force Attack	Ανεύρεση των παραμέτρων και του μυστικού κλειδιού με πολλαπλές δοκιμές τιμών με ή χωρίς τη χρήση λεξικού.
18	Mix Attack	Διαδοχικές επιθέσεις των παραπάνω μορφών.

Πίνακας 12: Η Αγγλική ορολογία των επιθέσεων σε υδατογραφημένα δεδομένα

5 Σχεδίαση σεναρίων υδατοσήμανσης και επιθέσεων

5.1 Η αρχιτεκτονική του συστήματος

Αυτή η ενότητα περιγράφει το σχεδιασμό του παραμετρικού λογισμικού στη δημιουργία και εκτέλεση σύνθετων σεναρίων και πειραμάτων που μπορούν να περιλαμβάνουν διαδοχικές υδατογραφήσεις, επιθέσεις και ανιχνεύσεις. Το λογισμικό έχει τη μορφή παραμετρικού προγράμματος που υλοποιήθηκε στη γλώσσα προγραμματισμού Java.

Εικόνα 12 : Η αρχιτεκτονική της δημιουργίας σεναρίων υδατογράφησης και επιθέσεων

Το σύστημα περιλαμβάνει τα παρακάτω δομικά στοιχεία.

ΑΑ	**Δομικό στοιχείο (module)**	**Περιγραφή**
1	Διασύνδεση	Αναλαμβάνει την ανάγνωση και γραφή συνόλων δεδομένων από ή σε αρχεία CSV.
2	Υδατογράφηση	Εκτελεί τη διαδικασία της δημιουργίας υδατογραφήματος στα δεδομένα με τη χρήση παραμέτρων και ενός μυστικού κλειδιού.
3	Ανίχνευση	Εκτελεί τη διαδικασία της αναζήτησης υδατογραφήματος σε δεδομένα με τη χρήση μυστικού κλειδιού. Το αποτέλεσμα επιβεβαιώνει (με ένα ποσοστό βεβαιότητας) τον εντοπισμό υδατογραφήματος στα δεδομένα που στη συνέχεια πιστοποιεί την ιδιοκτησία σε αυτόν που κατέχει το μυστικό κλειδί.
4	Επίθεση	Αναλαμβάνει τις γνωστές επιθέσεις εξάλειψης ή παραμόρφωσης του υδατογραφήματος.
5	Σενάριο	Χτίζει σενάρια ως γραμμικές αλληλουχίες υδατογραφήσεων και επιθέσεων με στόχο τη μελέτη της ανθεκτικότητας των σχημάτων και τη ρύθμιση των παραμέτρων για βέλτιστη αποτελεσματικότητα. Στη συνέχεια εκτελεί αυτά τα σενάρια δημιουργώντας ενδιάμεσες εκδόσεις των δεδομένων.
6	Αναφορά	Συλλέγει τα αρχεία καταγραφής που δημιουργούνται στα σενάρια ώστε να παρουσιάσουν τα αποτελέσματα μέσα από σχεδιαγράμματα στην εξαγωγή συμπερασμάτων.

Πίνακας 13 : Τα δομικά στοιχεία της δημιουργίας σεναρίων

Το σύστημα υλοποιήθηκε με το εργαλείο ανάπτυξης Eclipse IDE. Το πρόγραμμα είναι παραμετρικό, εκτελείται από γραμμή εντολών και μέσα δέσμες ενεργειών (scripts) με επαναληπτικές δομές, μπορεί να εκτελέσει σύνθετα σενάρια με διαφοροποιημένες παραμέτρους.

5.2 Ο αλγόριθμος της υδατογράφησης

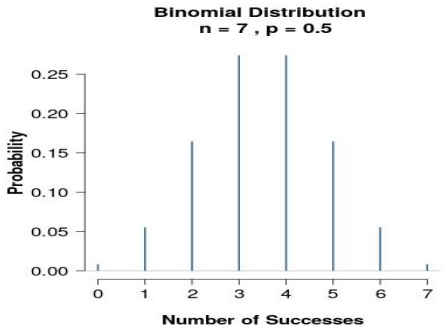
Ο αλγόριθμος που υλοποιήθηκε βασίζεται στο ανθεκτικό σχήμα του άρθρου [21] και περιλαμβάνει τη δημιουργία και τον εντοπισμό υδατογραφημάτων. Στηρίζεται σε παραμέτρους και σε μυστικό κλειδί, είναι σχήμα αόρατο, τυφλό, ισχυρό και αναστρέψιμο που βασίζεται στην αλλαγή τυχαία επιλεγμένων bits αριθμητικών τιμών σε τυχαίες πλειάδες και χαρακτηριστικά. Ο αλγόριθμος έχει υλοποιηθεί με τρόπο ώστε να αγνοεί κενές τιμές ή τιμές των οποίων το μέγεθος θα οδηγήσει σε απώλεια μεγάλου ποσοστού της πληροφορίας. Επίσης χαρακτηριστικά με ΜΗ αριθμητικές τιμές αγνοούνται με βάση τον έλεγχο τύπων δεδομένων που διαθέτει η Java. Οι μεταβλητές του αλγόριθμου όπως αναφέρονται στο [21] εξηγούνται στον επόμενο πίνακα.

Η προσθήκη και η ανίχνευση υλοποιήθηκαν με βάση τον ψευδοκώδικα του Πίνακα 14 που δέχεται ως παραμέτρους τις τιμές γ, ν, ξ, Κ (κλειδί), και α (όριο αποδοχής) με δεδομένες τις υποψήφιες αριθμητικές στήλες (A_i) για υδατοσήμανση που αποτελούν τα στοιχεία εισόδου του προγράμματος.

Προσθήκη	**Ανίχνευση**
1) **foreach** tuple $r \in R$ **do** 2) seed $\mathcal{G}$ with $r.P$ concatenated with $\mathcal{K}$ 3) **if** (next($\mathcal{G}$) mod γ equals 0) **then** // mark this tuple 4) attribute_index i = next($\mathcal{G}$) mod ν // mark attribute A_i 5) bit_index j = next($\mathcal{G}$) mod ξ // mark j^{th} bit 6) $r.A_i$ = mark(next($\mathcal{G}$), $r.A_i$, j) 7) mark(random_number i, value v, bit_index j) **return** value 8) **if** (i is even) **then** 9) set the j^{th} least significant bit of v to 0 10) **else** 11) set the j^{th} least significant bit of v to 1 12) **return** v	1) totalcount = matchcount = 0 2) **foreach** tuple $s \in S$ **do** 3) seed $\mathcal{G}$ with $s.P$ concatenated with $\mathcal{K}$ 4) **if** (next($\mathcal{G}$) mod γ equals 0) **then** // tuple was marked 5) attribute_index i = next($\mathcal{G}$) mod ν // A_i was marked 6) bit_index j = next($\mathcal{G}$) mod ξ // j^{th} bit was marked 7) totalcount = totalcount + 1 8) matchcount = matchcount + match(next($\mathcal{G}$), $s.A_i$, j) 9) τ = threshold(totalcount, α) 10) **if** ((matchcount $< \tau$) **or** (matchcount $>$ totalcount $- \tau$)) **then** 11) *suspect piracy* 12) match(random_number i, value v, bit_index j) **return** integer 13) **if** (i is even) **then** 14) return 1 if j^{th} least significant bit of v is 0 else return 0 15) **else** 16) return 1 if j^{th} least significant bit of v is 1 else return 0
Η γραμμή 2 αρχικοποιεί μια γεννήτρια τυχαίας ακολουθίας G με βάση το κλειδί της πλειάδας και το μυστικό κλειδί K. Έτσι, σε κάθε πλειάδα δημιουργείται ανεξάρτητη και μυστική ακολουθία αριθμών (με βάση το κλειδί) καθιστώντας το υδατογράφημα ανθεκτικό σε διαγραφές και ταξινομήσεις πλειάδων. Η ακολουθία χρησιμοποιείται για να αποφασιστεί αν η πλειάδα θα υδατοσημανθεί και σε αυτή την περίπτωση, η ακολουθία χρησιμοποιείται και για να καθορίσει το χαρακτηριστικό, τη θέση και την τιμή που θα δοθεί στο bit. Η γραμμή 3 καθορίζει άν η τρέχουσα πλειάδα θα επισημανθεί. Η γραμμή 4 καθορίζει το χαρακτηριστικό που θα επισημανθεί ανάμεσα στα ν υποψήφια. Η γραμμή 5 καθορίζει τη θέση του bit που θα επισημανθεί ανάμεσα στα ξ λιγότερο σημαντικά bit. Η συνάρτηση mark() [γραμμές 7-12] θέτει το επιλεγμένο bit σε 0 ή 1 σε σχέση με το αν ο τυχαίος αριθμός είναι άρτιος ή περιττός. Με αυτή τη λογική κάποιες τιμές αυξάνονται, κάποιες μειώνονται και κάποιες μένουν αμετάβλητες. Σε μια ενδεχόμενη επίθεση αφαίρεσης του υδατογραφήματος, ο επιτιθέμενος πρέπει να μαντέψει τις πλειάδες που έχουν σημαδευτεί, το χαρακτηριστικό μέσα σε κάθε πλειάδα, τη θέση bit και την τιμή που δόθηκε. Αυτός είναι ο λόγος που κάνει το υδατογράφημα τόσο ισχυρό. Χαρακτηριστικά με κενές τιμές παραλείπονται.	Η ανίχνευση προσδιορίζει τα bit που θα μπορούσαν να έχουν επισημανθεί στις γραμμές 1 έως 5 του αλγορίθμου εισαγωγής. Για κάθε αναγνωρισμένο bit, ο κώδικας ελέγχει εάν η τιμή του bit ταιριάζει με την τιμή που θα έπρεπε να έχει από τον αλγόριθμο εισαγωγής και μετράει τον αριθμό των ταιριασμάτων. Μεγάλος ή πολύ μικρός αριθμός οδηγεί σε υποψία πειρατείας. Η γραμμή 3 αρχικοποιεί μια γεννήτρια τυχαίας ακολουθίας G με βάση το κλειδί της πλειάδας και το μυστικό κλειδί, έτσι η ακολουθία των ψευδο-τυχαίων αριθμών θα είναι η ίδια με αυτή της αρχικής υδατοσήμανσης. Η γραμμή 4 καθορίζει εάν η τρέχουσα πλειάδα θα έπρεπε να έχει επιλεχθεί για σημάδι. Οι γραμμές 5 και 6 καθορίζουν το χαρακτηριστικό και το bit που θα έπρεπε να έχει επισημανθεί. Στη γραμμή 8 η συνάρτηση match() συγκρίνει την τρέχουσα τιμή bit με την τιμή που θα έπρεπε να έχει οριστεί για αυτό το bit στη δημιουργία. Η γραμμή 9 υπολογίζει πόσες πλειάδες δοκιμάστηκαν [totalcount] και πόσες περιέχουν την αναμενόμενη τιμή στα bits [matchcount]. Η συνάρτηση threshold() υπολογίζει το όριο τ. Η γραμμή 10 ελέγχει αν ο αριθμός [matchcount] είναι εντός αναμενόμενων ορίων.

Πίνακας 14: Επεξήγηση των αλγορίθμων προσθήκης και ανίχνευση υδατοσήμανσης

Ο παραπάνω αλγόριθμος υδατοσήμανσης καλεί τέσσερις φορές τη γεννήτρια τυχαίων αριθμών για κάθε επιλεγμένη πλειάδα. Η ανίχνευση αντιμετωπίζει την υδατοσήμανση ως επιλογή τιμών bits κατά αντιστοιχία. Αν η ανίχνευση ελέγξει ω bits, ο αριθμός των ταυτοποιήσεων ακολουθεί μια διωνυμική κατανομή με παραμέτρους ω και ½. Μια επιτυχής ανίχνευση αναμένει περίπου m=ω/2 ταυτοποιήσεις. Αν m<τ ή m>ω-τ δηλαδή είναι λίγες η πάρα πολλές οι ταυτοποιήσεις τότε δημιουργείται υπόνοια πειρατείας [21]. Έτσι ορίζονται ο δείκτης ακρίβειας WAR=match/total και ο δείκτης αλλοίωσης WDR=1-WAR. Το διάγραμμα 13 εμφανίζει μια κατανομή για ξ=7.



Εικόνα 13: Διωνυμική κατανομή ανίχνευσης για ξ=7 και πιθανότητα p=0,5

Υπάρχουν τρείς παραλλαγές του αλγόριθμου [21] που υλοποιήθηκαν στην εργασία και αφορούν:

i. Τη διαφοροποίηση στην τυχαία επιλογή των υποψήφιων χαρακτηριστικών (γραμμή 4) όπου με βάση βαρύτητες, οι πιθανότητες επιλογής χαρακτηριστικών διαφοροποιούνται δίνοντας περισσότερη έμφαση σε κάποια χαρακτηριστικά. Αν έχουμε ν χαρακτηριστικά Α, τότε θα πρέπει να ισχύει $\sum p_{Ai} = 1$, όπου p_{Ai} είναι η πιθανότητα επιλογής του χαρακτηριστικού Ai. Αν δοθούν οι βαρύτητες w1,w2,w3,w4,w5 που αντιστοιχούν σε 5 αριθμητικά χαρακτηριστικά, ο αλγόριθμος βρίσκει το άθροισμα ws=w1+w2+w3+w4+w5 και υπολογίζει την τιμή wR=ws*Rand(), όπου wR ϵ (0,ws]. Αν για παράδειγμα το wR ϵ (w1,w1+w2] τότε επιλέγεται η 2^{η} ιδιότητα, αν wR ϵ (w1+w2+w3,w1+w2+w3+w4], τότε επιλέγεται η 4^{η} ιδιότητα.

ii. Τη διαφοροποίηση στον αριθμό των LSBs σε διαφορετικά χαρακτηριστικά, με τη λογική ότι σε κάποια χαρακτηριστικά οι αριθμοί έχουν διαφορετική τάξη μεγέθους από άλλα. Έτσι στους μικρότερους αριθμούς θα έχουμε ελεγχόμενη και μικρότερη αναλογικά αλλοίωση των δεδομένων. Σε κάθε χαρακτηριστικό Ai, αποδίδεται διαφορετικό ξ_{Ai}.

iii. Τη χρήση συγκεκριμένων bits (PKB) ως εικονικό πρωτεύον κλειδί. Ο αλγόριθμος αντιμετωπίζει τις περιπτώσεις απουσίας κλειδιού όπου μπορεί να *χρησιμοποιηθεί το χαρακτηριστικό με τις λιγότερες διπλές τιμές* στο σύνολο του dataset. Ακόμα και με ένα χαρακτηριστικό, προτείνεται η χρήση συγκεκριμένων θέσεων bits για να παίξουν το ρόλο του εικονικού primary key και η χρήση άλλων bits για τη σήμανση. Για παράδειγμα ένα χαρακτηριστικό 20bit θα μπορούσε να αναλυθεί σε τρία μέρη : LSBs[0-5], PKBs[6-13] , MSBs[14-19], με στόχο την υλοποίηση σήμανσης με ένα μοναδικό χαρακτηριστικό.

Με βάση τον παραπάνω αλγόριθμο [21], στο πλαίσιο ανάλυσης της ανθεκτικότητας, αναφέρονται κάποιες υποθέσεις και συμπεράσματα που επιβεβαιώθηκαν και από τα πειράματα της εργασίας.

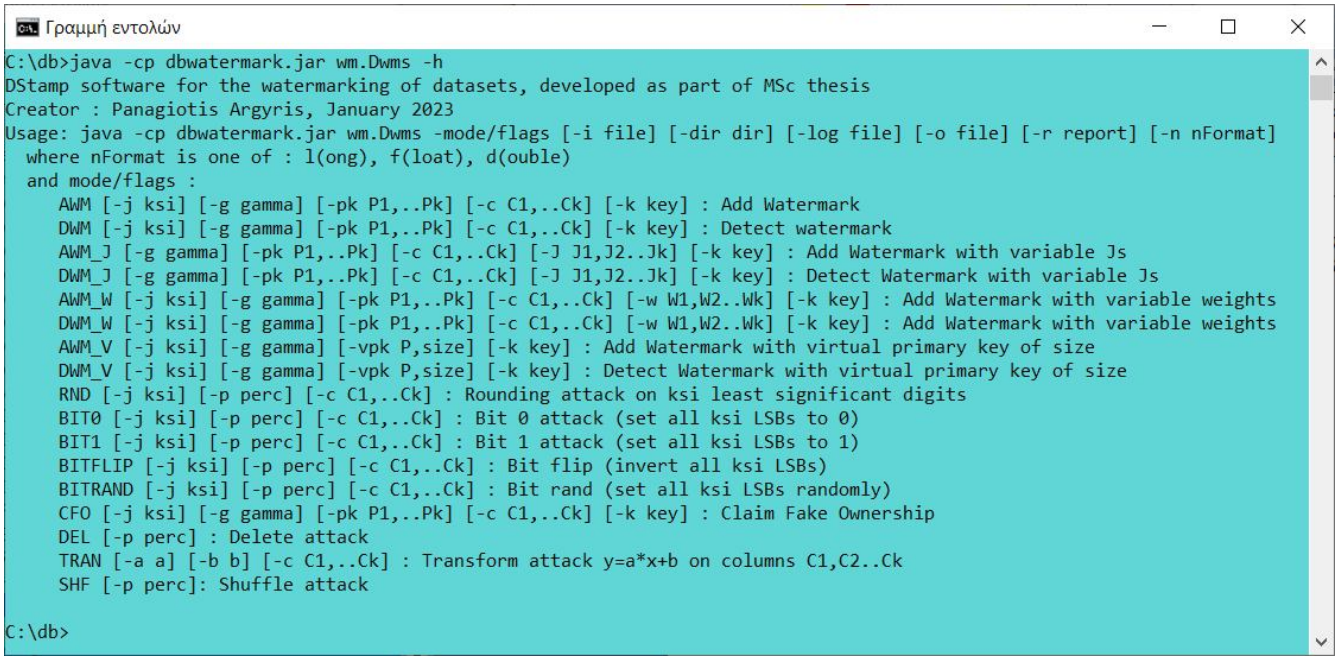
- Όσο μεγαλύτερη είναι η βάση δεδομένων τόσο ευκολότερη είναι η προστασία της.

- Ο αριθμός των πλειάδων που πρέπει να υδατοσημανθούν για δεδομένο βαθμό προστασίας δεν εξαρτάται από το μέγεθος της βάσης.
- Η καθορισμένη αναστροφή bits σε μια επίθεση (που αλλάζει το 0 σε 1 και το 1 σε 0) έχει περισσότερες πιθανότητες επιτυχίας από την τυχαία. Η αναστροφή bits σε όλες τις ξ θέσεις και σε όλες τις ν ιδιότητες επιφέρει αλλοίωση μεγέθους AM=M*ξ*ν/γ που είναι πολλαπλάσια (κατά ν) από την αλλοίωση του υδατογραφήματος (AY=M*ξ/γ).
- Αν ο επιτιθέμενος υπερεκτιμήσει το ξ, εισάγει μεγάλο σφάλμα χωρίς να αυξάνει την επιτυχία της επίθεσης, επίσης αν υποεκτιμήσει το ξ η επιτυχία θα είναι και πάλι μικρή.
- Η χρήση περισσότερων χαρακτηριστικών στο υδατόσημο δεν αλλάζει το μέγεθος της αλλοίωσης αλλά διαχέει το υδατόσημο καλύτερα.
- Υπάρχει μια σχέση ανάμεσα στο ξ και στο γ. Αν αυξηθεί το γ (μειωθεί το % των σημαδεμένων πλειάδων), μπορεί να αυξηθεί το ξ κρατώντας σταθερή τη συνολική αλλοίωση (AY=πλειάδες*ξ/γ). Η αύξηση των ξ και ν αυξάνουν την ανθεκτικότητα.

Η ανανέωση του υδατογραφήματος μετά από αλλαγές στη βάση, γίνεται εύκολα σε κάθε πλειάδα που αλλάζει [21], αφού η σήμανση κάθε πλειάδας εξαρτάται μόνο από το χαρακτηριστικό του πρωτεύοντος κλειδιού. Διαγραφή, εισαγωγή ή ενημέρωση μιας πλειάδας, δεν επηρεάζει τις υπόλοιπες πλειάδες. Αν αλλάξει η τιμή του πρωτεύοντος κλειδιού μιας πλειάδας, πρέπει να υπολογισθεί ξανά η σήμανση της πλειάδας. Επίσης αν ενημερωθεί οιοδήποτε άλλο χαρακτηριστικό και ο αλγόριθμος δεν το έχει επιλέξει δεν χρειάζεται άλλη ενέργεια. Διαφορετικά, αν ο αλγόριθμος το έχει επιλέξει, η σήμανση πρέπει να εφαρμοσθεί στο χαρακτηριστικό αυτό.

5.3 Το παραμετρικό λογισμικό υδατοσήμανσης και επιθέσεων

Το λογισμικό εκτελείται από γραμμή εντολών ενώ οι παράμετροι λειτουργίας του φαίνονται εκτελώντας το με την παράμετρο –h. Το παραμετρικό πρόγραμμα σχεδιάστηκε για να καλείται με ευελιξία μέσα από δέσμες ενεργειών (scripts).



Εικόνα 14:Οι παράμετροι του λογισμικού

Το πρόγραμμα διενεργεί υδατοσήμανση, επιθέσεις και ανίχνευση υδατοσήμανσης ανάλογα με τις παραμέτρους του. Κάθε υδατοσήμανση ή επίθεση διαβάζει ένα dataset ως είσοδο και παράγει ένα νέο dataset στην έξοδο με τις αλλαγές που επέφερε η υδατοσήμανση ή επίθεση. Στην ανίχνευση εμφανίζονται πληροφορίες στην κονσόλα ενώ ο βαθμός ανίχνευσης καταγράφεται σε αρχείο log.

Οι παράμετροι αφορούν :

-h : προβολή των παραμέτρων
-v : προβολή της έκδοσης
-i file : όνομα αρχείου εισόδου
-dir dirname : όνομα φακέλου εργασίας
-log file : όνομα αρχείου καταγραφής αποτελεσμάτων ανίχνευσης
-o file : όνομα αρχείου εξόδου
-r file : όνομα αρχείου καταγραφής μεταβολών τιμών
-n l(ong), **f**(loat), **d**(ouble) : η μορφή των δεδομένων.

Επειδή ο αλγόριθμος παρεμβαίνει σε bits των τιμών, πρέπει να καθορισθεί η μορφή και η ακρίβεια των αριθμητικών δεδομένων. Οι ακέραιοι αντιμετωπίζονται ως Long (64bit), οι πραγματικοί αριθμοί μικρής ακρίβειας ως αριθμοί κινητής υποδιαστολής Float(32 bit με 23 bits για τα δεκαδικά στοιχεία) ενώ οι πραγματικοί αριθμοί μεγάλης ακρίβειας ως Double(64 bit με 52 bits για τα δεκαδικά στοιχεία). Στα Float και Double, η παρέμβαση γίνεται στα LSBs της αριθμητικής τιμής (mantissa). Επειδή το ξ παρεμβαίνει στα LSBs, ένα μικρό ξ σε αριθμό μεγάλης ακρίβειας θα αλλοιώσει ελάχιστα τον αριθμό, αυτό όμως μπορεί να προδώσει την υδατογράφηση γιατί θα δημιουργήσει πολλά δεκαδικά εκεί που δεν υπήρχαν. Για παράδειγμα σε ένα χαρακτηριστικό που οι αριθμοί είναι της μορφής NNN.NNNN (πχ 167.2647) εάν αλλάξουμε ένα-δύο λιγότερα σημαντικά LSBs ο αριθμός μπορεί να καταλήξει 167.2646000000001 που διαφέρει σε λεπτομέρεια από τις τιμές χαρακτηριστικών άλλων πλειάδων και μπορεί να προδώσει πληροφορίες της υδατογράφησης, τα επιλεγμένα χαρακτηριστικά, τη συχνότητα των πλειάδων (γ) και το βαθμό παρέμβασης (ξ). Ο τύπος δεδομένων BigDecimal της Java αν και ακριβής στην αναπαράσταση μεγάλων δεκαδικών είναι πιο αργός στους υπολογισμούς και δεν κρίθηκε κατάλληλος για τη διαχείριση των bits.

-AWM [-j ksi] [-g gamma] [-pk P1,..Pk] [-c C1,..Ck] [-k key] : Προσθήκη υδατογραφήματος με ξ, γ, μυστικό κλειδί key, στις στήλες με αριθμούς C1,..Ck και πρωτεύον κλειδί τις στήλες P1,..Pk

- DWM [-j ksi] [-g gamma] [-pk P1,..Pk] [-c C1,..Ck] [-k key] : Ανίχνευση υδατογραφήματος με ξ, γ, μυστικό κλειδί key, στις στήλες με αριθμούς C1,..Ck και πρωτεύον κλειδί τις στήλες P1,..Pk

- AWM_J [-g gamma] [-pk P1,..Pk] [-c C1,..Ck] [-J J1,.Jk] [-k key] : $1^{η}$ παραλλαγή προσθήκης υδατογραφήματος με γ, μυστικό κλειδί key, στήλες με αριθμούς C1,..Ck, πρωτεύον κλειδί τις στήλες P1,..Pk και αντίστοιχα ξ ανά στήλη J1,J2..Jk

- DWM_J [-g gamma] [-pk P1,..Pk] [-c C1,C2..Ck] [-J J1,.Jk] [-k key] : $1^{η}$ παραλλαγή ανίχνευσης υδατογραφήματος με γ, μυστικό κλειδί key, στήλες με αριθμούς C1,..Ck, πρωτεύον κλειδί τις στήλες P1,..Pk και αντίστοιχα ξ ανά στήλη J1,J2..Jk

- AWM_W [-j ksi] [-g gamma] [-pk P1,..Pk] [-c C1,..Ck] [-w W1,..Wk] [-k key] : $2^{η}$ παραλλαγή προσθήκης υδατογραφήματος με παραμέτρους γ, μυστικό κλειδί key, στις στήλες με αριθμούς C1,C2..Ck, πρωτεύον κλειδί τις στήλες P1,..Pk και πιθανότητα επιλογής τους W1,W2..Wk

- DWM_W [-j ksi] [-g gamma] [-pk P1,..Pk] [-c C1,..Ck] [-w W1,..Wk] [-k key] : 2^{η} παραλλαγή ανίχνευσης υδατογραφήματος με παραμέτρους γ, μυστικό κλειδί key, στις στήλες με αριθμούς C1,C2..Ck, πρωτεύον κλειδί τις στήλες P1,..Pk και πιθανότητα επιλογής τους W1,W2..Wk

- AWM_V [-j ksi] [-g gamma] [-vpk P,size] [-k key] : 3^{η} παραλλαγή προσθήκης υδατογραφήματος με ξ, γ, εικονικό πρωτεύον κλειδί στήλης P, μεγέθους size , μυστικό κλειδί key.

- DWM_V [-j ksi] [-g gamma] [-vpk P,size] [-k key] : 3^{η} παραλλαγή ανίχνευσης υδατογραφήματος με ξ, γ, εικονικό πρωτεύον κλειδί στήλης P, μεγέθους size , μυστικό κλειδί key.

- RND [-j ksi] [-p perc] [-c C1,C2..Ck] : Επίθεση στρογγυλοποίησης σε ξ δεκαδικά ψηφία, στις στήλες C1,C2..Ck και σε perc ποσοστό επίθεσης (πλειάδων)

- BIT0 [-j ksi] [-p perc] [-c C1,C2..Ck] : Επίθεσης δυαδικού ψηφίου, όπου όλα τα ξ LSB ψηφία μετατρέπονται σε 0, στις στήλες C1,C2..Ck και σε perc ποσοστό επίθεσης (πλειάδων)

- BIT1 [-j ksi] [-p perc] [-c C1,C2..Ck] : Επίθεσης δυαδικού ψηφίου, όπου όλα τα ξ LSB ψηφία μετατρέπονται σε 1, στις στήλες C1,C2..Ck και σε perc ποσοστό επίθεσης (πλειάδων)

- BITFLIP [-j ksi] [-p perc] [-c C1,C2..Ck] : Επίθεσης δυαδικού ψηφίου, όπου όλα τα ξ LSB ψηφία αντιστρέφονται, στις στήλες C1,C2..Ck και σε perc ποσοστό επίθεσης (πλειάδων)

- BITRAND [-j ksi] [-p perc] [-c C1,C2..Ck] : Επίθεσης δυαδικού ψηφίου, όπου όλα τα ξ LSB ψηφία παίρνουν τυχαίες τιμές, στις στήλες C1,C2..Ck και σε perc ποσοστό επίθεσης (πλειάδων)

- CFO [-j ksi] [-g gamma] [-c C1,C2..Ck] [-k key] : Προσθήκη ΨΕΥΤΙΚΟΥ υδατογραφήματος με παραμέτρους ξ, γ, μυστικό κλειδί στις στήλες του dataset με αριθμούς C1,C2..Ck

- DEL [-p perc] : Επίθεση διαγραφής υποσυνόλου σε perc ποσοστό επίθεσης (πλειάδων)

- TRAN [-a a] [-b b] [-c C1,C2..Ck] : Επίθεση γραμμικού μετασχηματισμού τιμών με βάση τη σχέση y=a*x+b στις στήλες C1,C2..Ck

- SHF [-p perc]: Επίθεση αντιμετάθεσης γειτονικών πλειάδων σε perc ποσοστό επίθεσης (πλειάδων)

Για την αποτελεσματική παραμετροποίηση μιας υδατογράφησης αλλά και μιας επιτυχημένης επίθεσης πρέπει να προηγηθεί μια επισκόπηση των δεδομένων, ώστε :

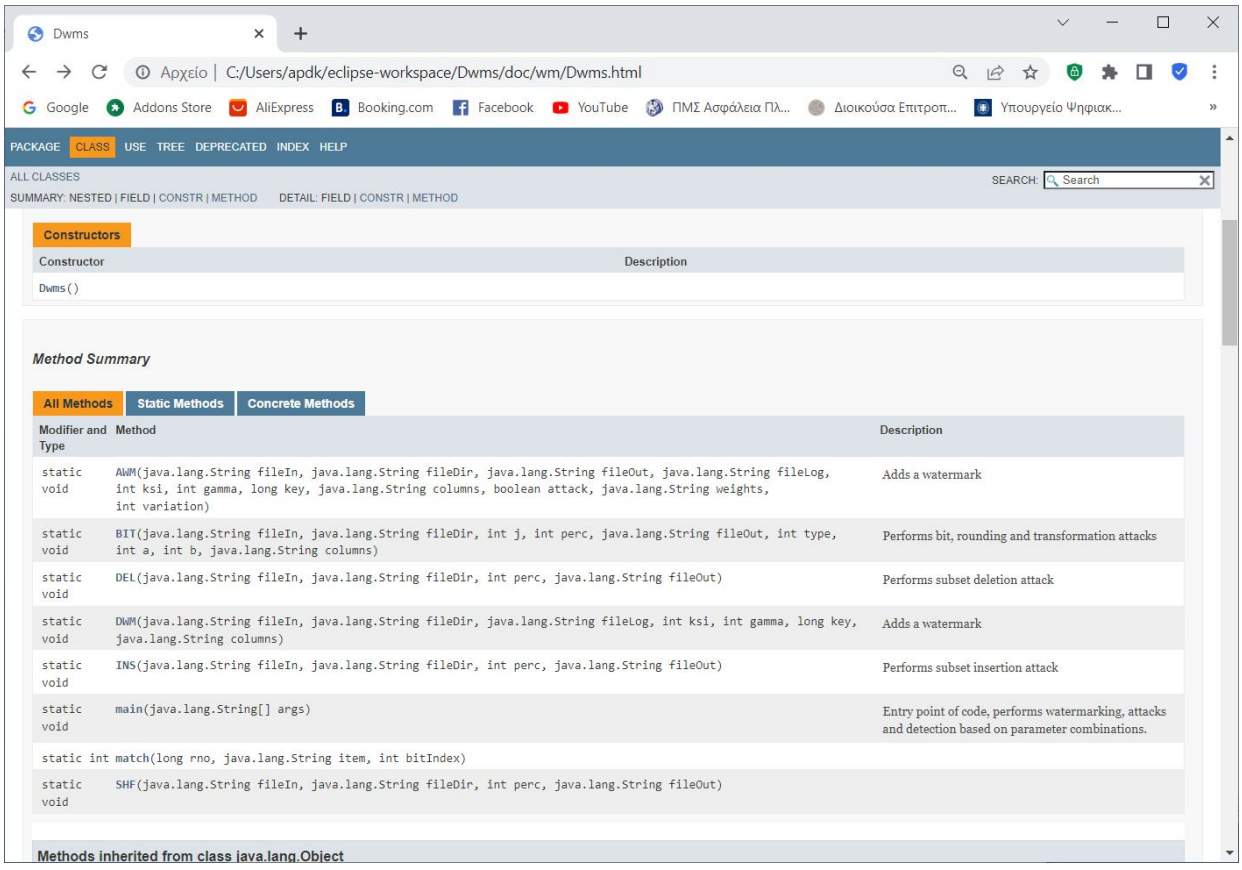
- Να επιλεχθούν τα χαρακτηριστικά κλειδιά των πλειάδων ή αυτά που μπορούν να παίξουν το ρόλο κλειδιών [-pk P1,..Pk] ή εικονικών κλειδιών [-vpk size]
- Να επιλεχθούν οι στήλες που περιλαμβάνουν αριθμητικά στοιχεία, το εύρος των τιμών, η ύπαρξη πολλών κενών τιμών [-c C1,C2..Ck]
- Να επιλεχθεί η μορφή των αριθμών που θα καθορίσει την επιλογή του τύπου -n (Long, Float, ή Double)
- Να επιλεχθεί το κατάλληλο ξ με βάση τα παραπάνω [-j ksi] που δεν θα αλλοιώσει τις τιμές ούτε στο περιεχόμενο ούτε στην εμφάνιση. Για παράδειγμα, ο αριθμός 254.26000021 μπορεί

να περιλαμβάνει μικρή αλλοίωση του 254.26 όμως προδίδει την ύπαρξη της αλλοίωσης από τη μορφή.

Αν δεν δοθεί αρχείο εξόδου (-o), το πρόγραμμα το δημιουργεί αυτόματα με αφετηρία το όνομα αρχείου εισόδου και προσθέτοντας ονοματα και τιμές παραμέτρων κατά την κλήση. Για παράδειγμα, ξεκινώντας με το αρχείο covid2022.csv, κάνοντας α) προσθήκη υδατογραφήματος, στη συνέχεια β) διαγραφή υποσυνόλου, στη συνέχεια γ) μετασχηματισμό και στη συνέχεια δ) προσθήκη ψεύτικου υδατογραφήματος, θα προκύψουν ενδιάμεσα αρχεία :

```
covid22.csv
covid22.csv.AWMj5g10c20,21,30,32key1073.csv
covid22.csv.AWMj5g10c20,21,30,32key1073.DELp25.csv
covid22.csv.AWMj5g10c20,21,30,32key1073.DELp25.TRANSa2b4.csv
covid22.csv.AWMj5g10c20,21,30,32key1073.DELp25.TRANSa2b4.CFOj5g8c20,32key459.csv
```

Το λογισμικό συνοδεύεται από τεκμηρίωση μορφής Javadoc.

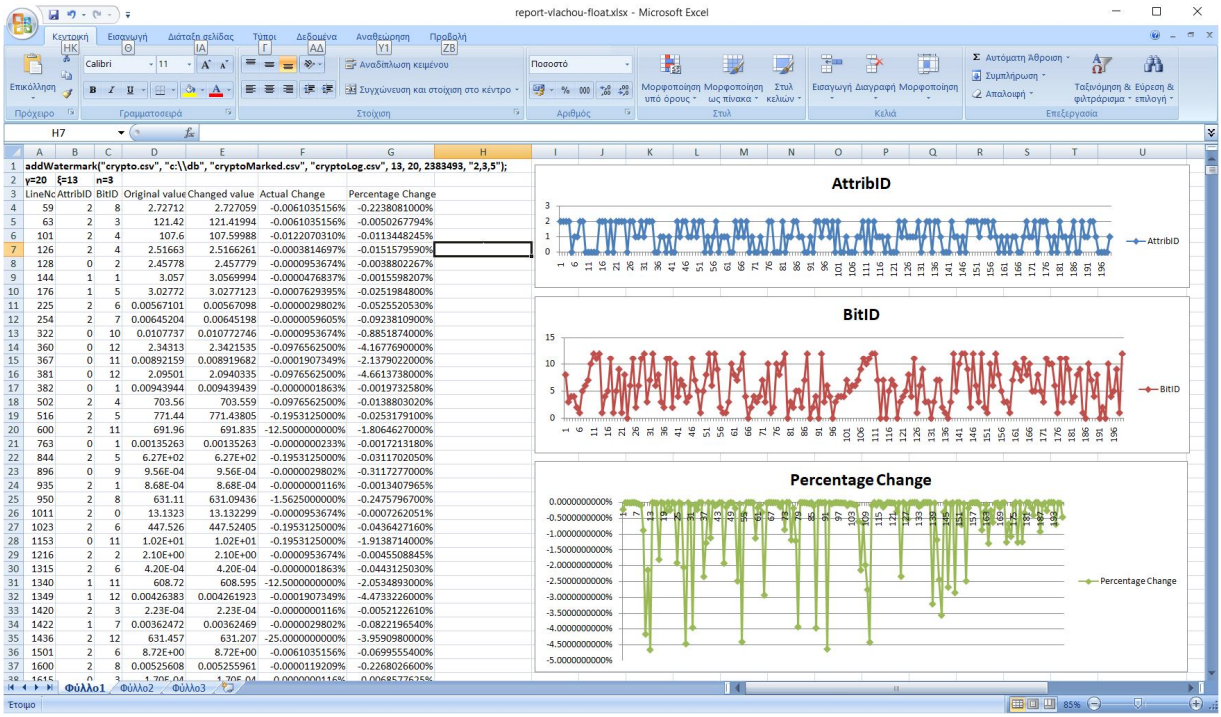


Εικόνα 15: Η τεκμηρίωση του λογισμικού σε τυποποιημένη μορφή Javadoc

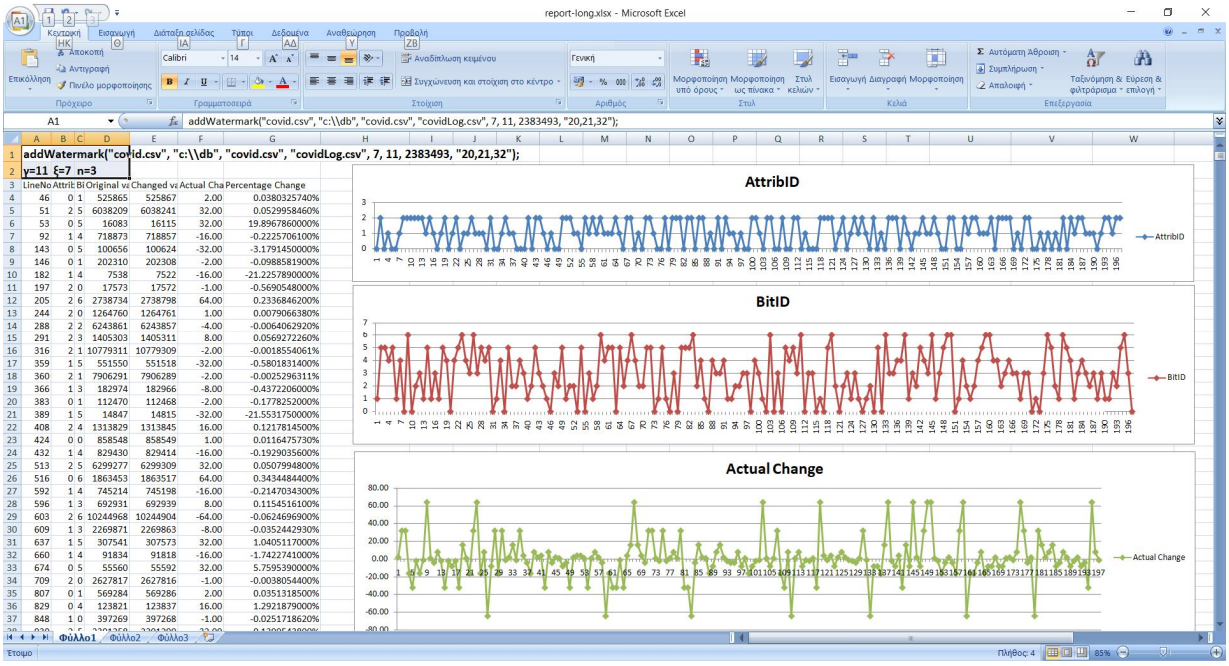
Στη φάση της ανάπτυξης, χρήσιμα στοιχεία των υδατογραφήσεων καταγράφτηκαν και αναλύθηκαν σε φύλλα εργασίας του MS Excel στην παραγωγή διαγραμμάτων με στόχο τον έλεγχο της υλοποίησης. Οι πληροφορίες για τον έλεγχο της υδατογράφησης αφορούσαν :

- Τον αριθμό της πλειάδας που άλλαξε
- Τον αριθμό του χαρακτηριστικού που επιλέχθηκε
- Τον αριθμό bit του LSB που επιλέχθηκε
- Την αρχική και τελική τιμή της τιμής του χαρακτηριστικού

- Την πραγματική και ποσοστιαία μεταβολή της τιμής



Εικόνα 16: Έλεγχος της λειτουργίας του λογισμικού (σε αριθμητικές στήλες τύπου float)

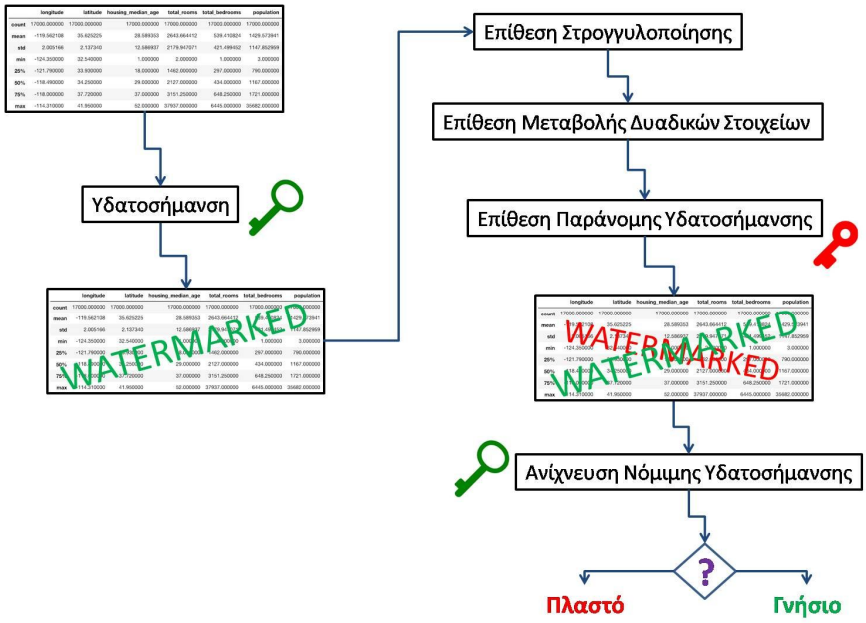


Εικόνα 17: Έλεγχος της λειτουργίας του λογισμικού (σε αριθμητικές στήλες τύπου long)

Το παράρτημα ΙΙΙ περιλαμβάνει διαγράμματα ελέγχου του λογισμικού (Εικόνες 37, 38, 43-46).

5.4 Η δημιουργία σύνθετων σεναρίων

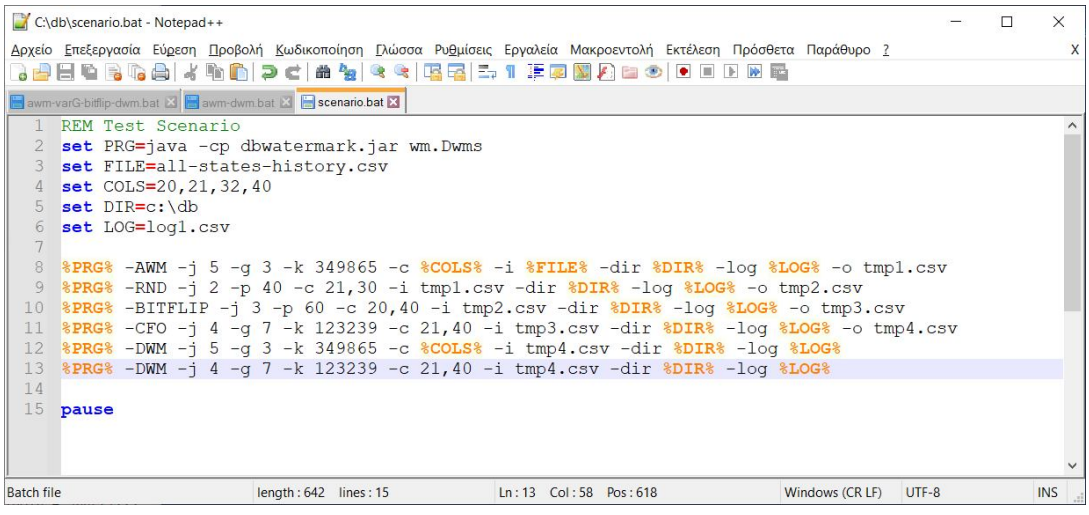
Με δεδομένο το παραμετρικό λογισμικό δίνεται η δυνατότητα δημιουργίας σεναρίων διατυπωμένων σε μορφή δεσμών ενεργειών (scripts) με επαναληπτικές υδατογραφήσεις, επιθέσεις και ανιχνεύσεις και διαφορετικούς συνδυασμούς παραμέτρων κάθε φορά.



Εικόνα 18: Παράδειγμα σεναρίου που περιλαμβάνει τρείς διαδοχικές επιθέσεις

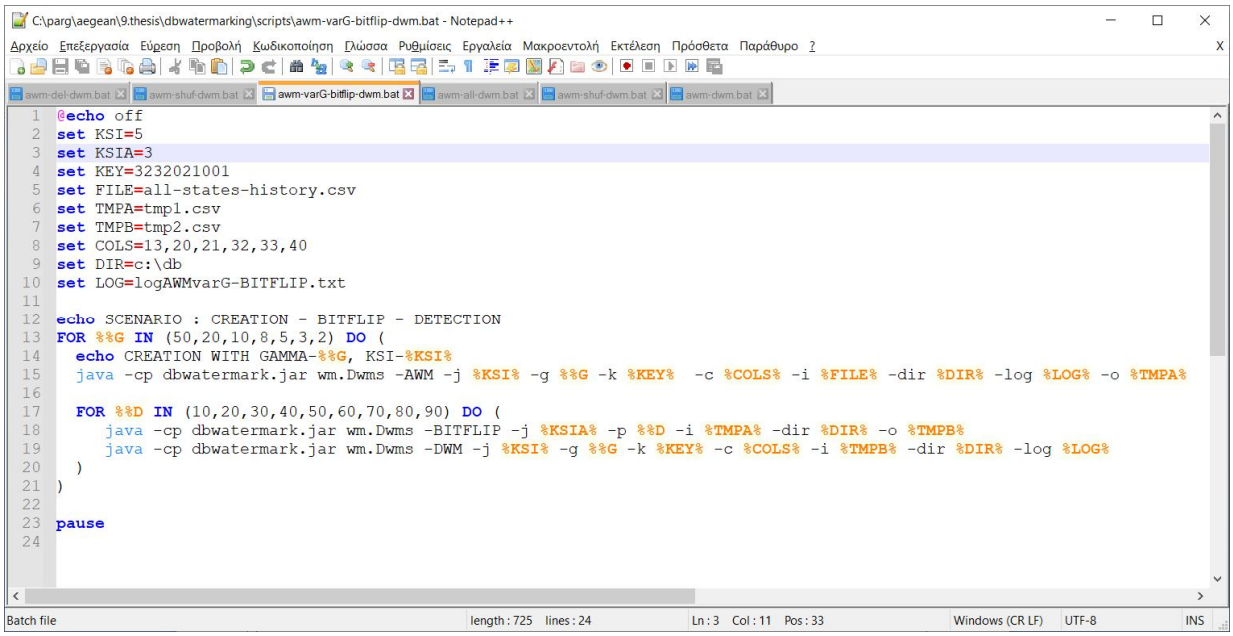
Ο στόχος των σεναρίων είναι να δοκιμάσουν συνδυασμούς υδατογραφήσεων και επιθέσεων ώστε να εντοπίσουν τους συνδυασμούς των παραμέτρων που οδηγούν σε υψηλή αποτελεσματικότητα με την ελάχιστη παραμόρφωση των δεδομένων. Επειδή αυτό απαιτεί την εκτέλεση σύνθετων σεναρίων αλλάζοντας τις παραμέτρους των επιμέρους ενεργειών, αναπτύχθηκε μια εύκολη διαδικασία διατύπωσης και εκτέλεσης σεναρίων με προσωρινή αποθήκευση των δεδομένων σε κάθε στάδιο της επεξεργασίας καθώς και ενιαίο log.

Το σενάριο με τις ενέργειες : Add WM => Rnd Attack => BitFlip Attack => Add Fake WM => Detect Real WM => Detect Fake WM έχει αποτυπωθεί ως εκτελέσιμο στην Εικόνα 19.



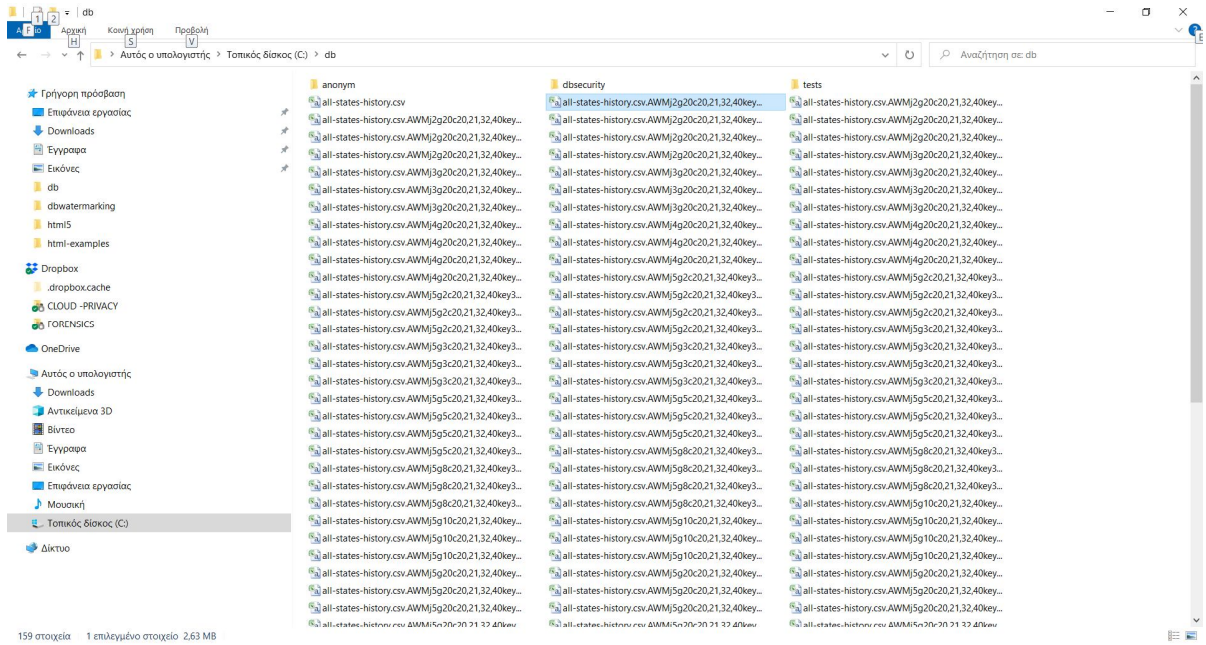
Εικόνα 19: Σενάριο υδατογραφήσεων και επιθέσεων σε μορφή δέσμης ενεργειών

Επίσης μπορούν να διατυπωθούν επαναληπτικά σενάρια υδατογραφήσεων και επιθέσεων με συνδυασμούς παραμέτρων ώστε να βρεθεί ο συνδυασμός των παραμέτρων που οδηγεί σε ανθεκτικότερα σχήματα. Εκτέλεση σεναρίων με υδατογραφήσεις γ=50,20,10,8,5,3,2 και επιθέσεις BITFLIP ποσοστών αλλοίωσης 10 έως 90% (7*9=63 συνδυασμοί) φαίνεται στην Εικόνα 20.



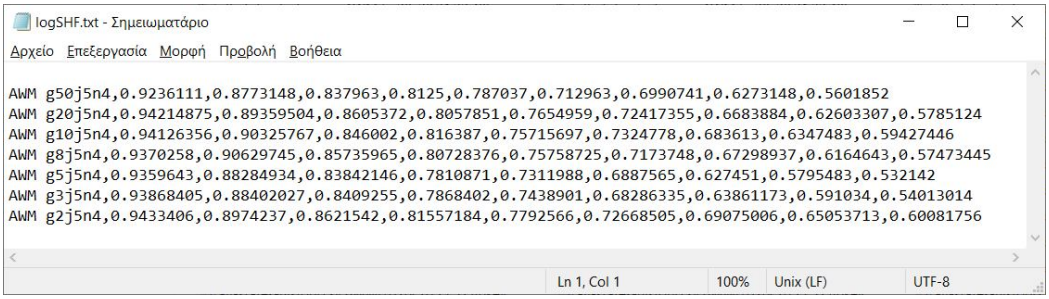
Εικόνα 20: Επαναληπτική εκτέλεση ενός σεναρίου με συνδυασμούς παραμέτρων

Αν στα scripts δεν δίνονται ονόματα για τα ενδιάμεσα αρχεία εξόδου, θα δημιουργηθούν πολλαπλά αρχεία των ενδιάμεσων σταδίων στο δίσκο για μελλοντικό έλεγχο.

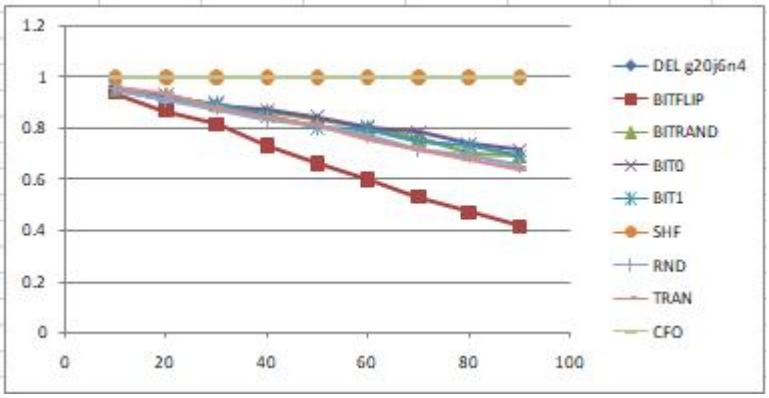


Εικόνα 21: Προσωρινά αρχεία επαναληπτικών σεναρίων

Τα αποτελέσματα των επαναληπτικών εκτελέσεων καταγράφονται σε ενιαίο αρχείο για εισαγωγή στο MS Excel και τη δημιουργία γραφημάτων όπως δείχνουν οι Εικόνες 22 και 23.



Εικόνα 22: Αποτελέσματα ανιχνεύσεων με συνδυασμούς παραμέτρων δημιουργίας και επιθέσεων



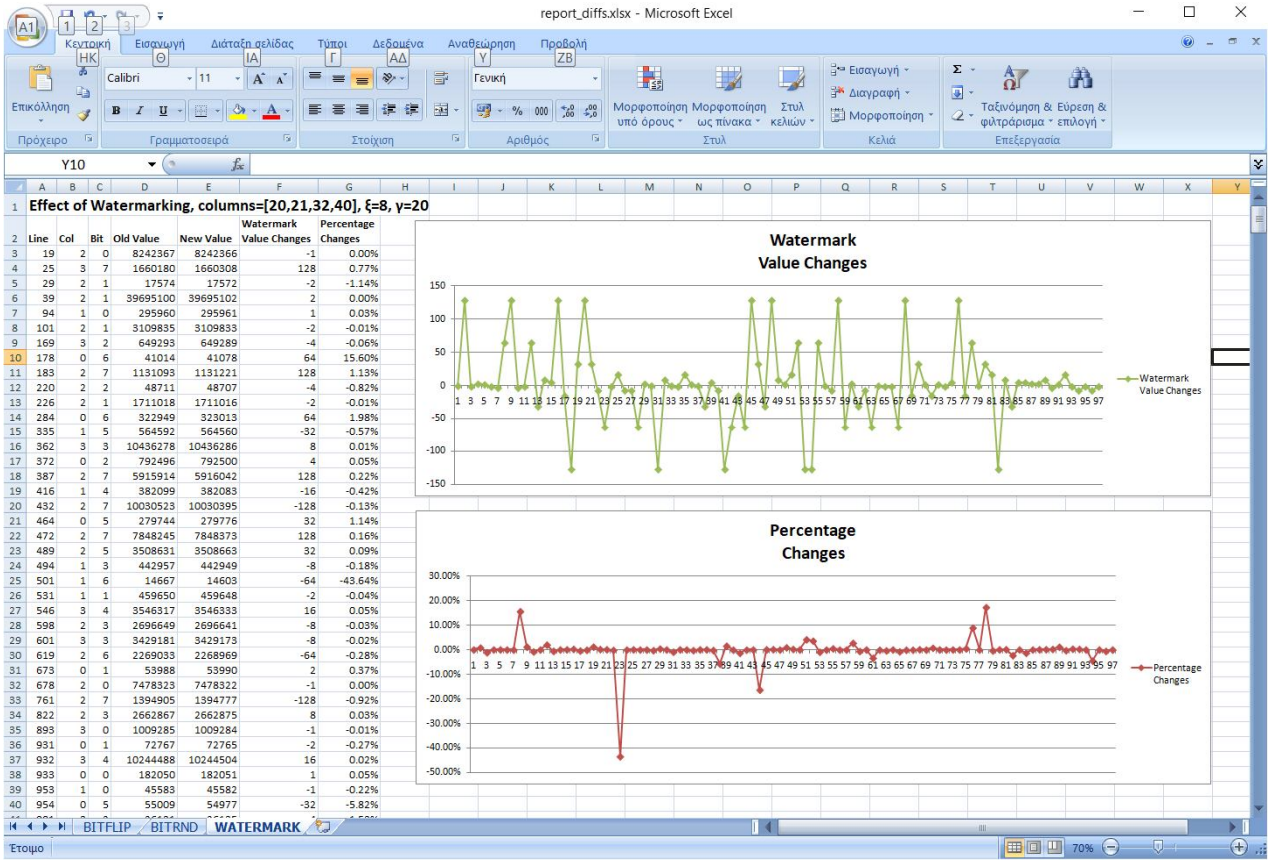
Εικόνα 23: Η μορφή των συγκεντρωτικών γραφημάτων που αποτυπώνουν την επιτυχία της ανίχνευσης στον ψ άξονα, μέσα από σενάρια διαφορετικών επιθέσεων με μεταβλητό το ποσοστό επίθεσης στον x-άξονα

Ένας εναλλακτικός τρόπος που χρησιμοποιήθηκε στα αρχικά στάδια των πειραμάτων στη δημιουργία των κλήσεων του παραμετρικού προγράμματος, ήταν η χρήση φύλλων εργασίας με υπολογιζόμενα κελιά με παραμέτρους για τη μορφοποίηση της απαιτούμενης εντολής στην τελευταία στήλη.

AA	Ενέργεια	Περιγραφή	ξ	γ	Επιθ (%)	Στήλες	Κλειδί	Είσοδος	Έξοδος	Εντολή
1	AWM	Add WM	5	3		20,32,40	3248349	covid.csv	tmp1.csv	java -cp dbwatermark.jar wm.Dwms -AWM -j 5 -g 3 -k 3248349 -c 20,32,40 -i covid.csv -o tmp1.csv
2	RND	Rnd Attack	6		40	32,40,42		tmp1.csv	tmp2.csv	java -cp dbwatermark.jar wm.Dwms -RND -p 40 -j 6 -k 32,40,42 -i tmp1.csv -o tmp2.csv
3	BIT0	Bit0 Attack	3		50	32,40,42		tmp2.csv	tmp3.csv	java -cp dbwatermark.jar wm.Dwms -BIT0 -p 50 -j 3 -k 32,40,42 -i tmp2.csv -o tmp3.csv
4	CFO	Fake Ownership	4	7		32,40,41	123239	tmp3.csv	tmp4.csv	java -cp dbwatermark.jar wm.Dwms -CFO -j 4 -g 7 -k 123239 -c 32,40,41 -i tmp3.csv -o tmp4.csv
5	DWM	Detect Legal	5	3		20,32,40	3248349	tmp4.csv		java -cp dbwatermark.jar wm.Dwms -DWM -j 5 -g 3 -c 20,32,40 -k 3248349 -i
6	DEL	Deletion Attack	20	0	30			tmp4.csv	tmp5.csv	java -cp dbwatermark.jar wm.Dwms -DEL -p 30 -i tmp4.csv -o tmp5.csv
9	DWM	Detect Legal	5	3		20,32,40	3248349	tmp5.csv		java -cp dbwatermark.jar wm.Dwms -DWM -j 5 -g 3 -c 20,32,40 -k 3248349 -i

Πίνακας 15: Χρήση MS Excel στη δημιουργία παραμετρικών κλήσεων

Με τη βοήθεια του λογισμικού, που καταγράφει τις μεταβολές και τις ποσοστιαίες μεταβολές των τιμών (με την παράμετρο –r), δημιουργήθηκαν οπτικές εικόνες των αλλαγών που επιφέρει ένα υδατογράφημα στα δεδομένα φορτώνοντας τα δεδομένα σε MS Excel.



Εικόνα 24: Μεταβολές που επιφέρει ένα υδατογράφημα με τιμές ξ=8 και γ=20

Συνοψίζονται, οι καινοτομίες στην εκπόνηση των πειραμάτων μέσω του λογισμικού :

I. Η υλοποίηση ενός εύκολου στη χρήση παραμετρικού λογισμικού υδατογράφησης, επιθέσεων και ανιχνεύσεων που επεξεργάζεται άμεσα αρχεία με σύνολα δεδομένων (CSV) χωρίς την απαίτηση βάσης δεδομένων.

II. Η ανάπτυξη και εκτέλεση σύνθετων και επαναληπτικών σεναρίων επιθέσεων με αυτοματοποιημένο τρόπο μέσα από εκτελέσιμες δέσμες ενεργειών.

III. Η εκπόνηση πολλαπλών πειραμάτων με διαφοροποιημένες παραμέτρους, η παραγωγή και οργάνωση όλων των αποτελεσμάτων και γραφημάτων.

6 Πειράματα ανθεκτικότητας υδατογραφήσεων

6.1 Η υποδομή σε εξοπλισμό Η/Υ των πειραμάτων

Τα χαρακτηριστικά του Η/Υ που εκτελέστηκαν τα πειράματα είναι : επεξεργαστής Intel Core i7, δίσκος 500GB SSD, 16GB μνήμη RAM, λειτουργικό σύστημα Windows 10 Pro 64 bit. Η εκτέλεση του συνόλου των σεναρίων σε 3 datasets απαίτησε 70 λεπτά.

6.2 Τα σύνολα δεδομένων

Τα σύνολα δεδομένων με τα οποία έγιναν τα πειράματα αντλήθηκαν από την πηγή https://www.kaggle.com/datasets, αφορούν αριθμητικά δεδομένα που διαφέρουν στο πλήθος των πλειάδων, στο πλήθος των χαρακτηριστικών, στα επιλεγμένα πρωτεύοντα κλειδιά και στους τύπους των αριθμητικών δεδομένων (Long, Float). Τα datasets των πειραμάτων είναι :

- Περιστατικά covid USA με 20780 πλειάδες, 42 χαρακτηριστικά (Long), σύνθετο κλειδί (ημ/νια + κωδικός χώρας), πηγή :

  https://www.kaggle.com/datasets/ramjasmaurya/covid-in-usa-states
- Διακυμάνσεις τιμών κρυπτονομισμάτων με 72946 πλειάδες, 6 χαρακτηριστικά (Float), σύνθετο κλειδί (νόμισμα+ημ/νία), πηγή :

  https://www.kaggle.com/datasets/maharshipandya/-cryptocurrency-historical-prices-dataset
- Διακυμάνσεις 33 νομισμάτων με 6510 πλειάδες, 120 χαρακτηριστικά (Float και Long), κλειδί (ημ/νία), πηγή :

  https://www.kaggle.com/datasets/mukhazarahmad/worldwide-stock-market-indices-data

6.3 Οι επιθέσεις που υλοποιήθηκαν

Με βάση την υλοποίηση του αλγορίθμου [21], η υδατογράφηση, οι επιθέσεις και η ανίχνευση αξιοποιούν σημαντικές παραμέτρους που μπορούν να ρυθμιστούν κατά περίπτωση:

i. γ, η παράμετρος που καθορίζει τη συχνότητα των επιλεγμένων πλειάδων (που οδηγεί σε 100/γ % επιλεγμένων πλειάδες)

ii. Η λίστα των χαρακτηριστικών για σήμανση/επίθεση σε μορφή $[c_1,c_2, \dots c_v]$ όπου c_i είναι οι θέσεις των στηλών. Για παράδειγμα σε ένα σύνολο με 40 χαρακτηριστικά, η λίστα [3,5,8,12] καθορίζει ότι οι επιθέσεις θα γίνουν στις στήλες με αριθμό 3,5,8 και 12. Αυτό καθορίζει το ν (=4), τον αριθμό των διαθέσιμων χαρακτηριστικών για σήμανση και για επιθέσεις.

iii. Τα χαρακτηριστικά με ρόλο πρωτεύοντος κλειδιού ή τα bits του εικονικού κλειδιού

iv. ξ, ο αριθμός των λιγότερο σημαντικών δυαδικών ψηφίων (LSB) για σήμανση (που οδηγεί σε 1 τυχαίο επιλεγμένο bit σε κάθε επιλεγμένο χαρακτηριστικό της επιλεγμένης πλειάδας)

v. κ, το μυστικό κλειδί της υδατοσήμανσης ή της ανίχνευσης

vi. Παραλλαγές που καθορίζουν διαφορετικά ξ στα χαρακτηριστικά ή διαφορετική συχνότητα επιλογής χαρακτηριστικών

Οι επιθέσεις που υλοποιήθηκαν συνολικά περιλαμβάνουν συνδυασμούς επτά τύπων επιθέσεων:

i. **Bit attack** με παραλλαγές Bit0, Bit1, BitFlip, BitRnd όπου τα bits των αριθμητικών δεδομένων μετατρέπονται σε 0 σε 1 ή αντιστρέφονται ή αλλάζουν τυχαία οι τιμές τους, σε **επιλεγμένες πλειάδες** (σε ποσοστό 100/γ %), **σε ξ LSBs** [0…ξ-1] και σε **επιλεγμένα αριθμητικά χαρακτηριστικά του dataset** (πλήθους ν). Η αλλοίωση είναι ξ*ν bits σε κάθε πλειάδα και 100*ξ*ν/γ % συνολικά για το dataset. Αυτό επιφέρει πολύ μεγαλύτερη ζημιά από το υδατογράφημα που επηρεάζει ένα τυχαίο χαρακτηριστικό σε κάθε τυχαία πλειάδα με συνολική αλλοίωση : 100*ξ/γ %.

ii. **Rounding attack**, επίθεση στρογγυλοποίησης που επιλέγει το βαθμό στρογγυλοποίησης με βάση έναν αριθμό δεκαδικών ψηφίων (παράμετρος ξ) **σε όλες τις πλειάδες** κάποιων επιλεγμένων αριθμητικών χαρακτηριστικών (για παράδειγμα των [3,5,8,12]).

iii. **Transformation**, επίθεση γραμμικού μετασχηματισμού που μετατρέπει αριθμητικά δεδομένα σε άλλη κλίμακα. Η υλοποίηση μετατρέπει αριθμητικές τιμές σε διαφορετικές μονάδες μέτρησης μέσα από σχέσεις Y=a*X+b. Για παράδειγμα : F=C*9/5 + 32 (θερμοκρασία), ft=0,3048m (απόσταση), lb(pound)=2,2046*Kg. Εδώ θα πρέπει κάποιος να δει τις μονάδες μέτρησης των χαρακτηριστικών και να κάνει στοχευμένες μετατροπές στα επιλεγμένα χαρακτηριστικά στο σύνολο των πλειάδων.

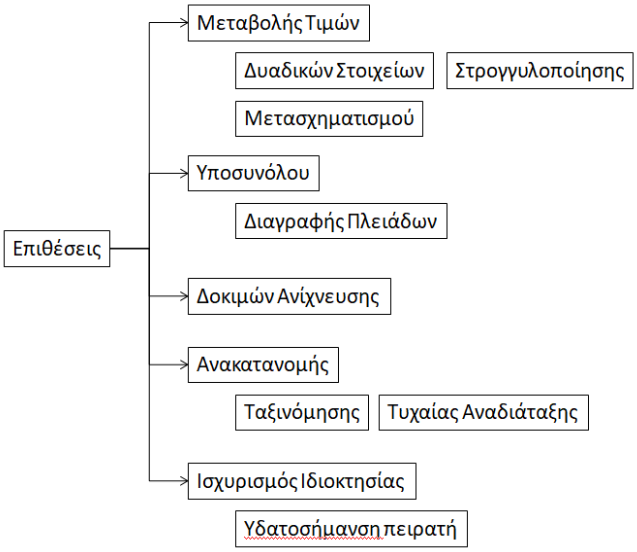
iv. **Subset elimination**, επιθέσεις όπου διαγράφονται πλειάδες σε τυχαίες θέσεις.

v. **Brute Force attack,** επιθέσεις δοκιμής κλειδιών και παραμέτρων

vi. **Tupple shuffle**, αντιμετάθεση ζευγών πλειάδων, στην υλοποίηση επιλέγονται τυχαία πλειάδες (με ποσοστό επίθεσης %) και αντιμετατίθενται με την επόμενη πλειάδα.

vii. **False claim of ownership**, ψεύτικη υποστήριξη ιδιοκτησίας με την εισαγωγή πλαστού υδατοσήματος. Αυτό υλοποιείται ως μια προσθήκη 2^{ου} υδατογραφήματος όπου οι παράμετροι, γ, ξ, ν και το κλειδί είναι γνωστά μόνο στον πειρατή, ώστε μόνο αυτός να μπορεί να επιβεβαιώσει το δικό του υδατογράφημα. Ένα παράνομο υδατόσημο έχει επιτυχία αν αλλοιώσει το αυθεντικό προβλέποντας τα γ, ξ και ν του δημιουργού. Αν το νόμιμο υδατόσημο δεν αλλοιωθεί από το παράνομο, συνυπάρχουν τα 2 υδατοσήματα και εποτελεί δύσκολη περίπτωση διάκρισης του νόμιμου ιδιοκτήτη.

Τα σενάρια που δοκιμάστηκαν περιλαμβάνουν ακολουθίες διαφορετικών επιθέσεων.



Εικόνα 25: Οι επιθέσεις που υλοποιήθηκαν στην εργασία

Οι επιθέσεις έγιναν με παραμετρικό τρόπο ελέγχοντας το βαθμό της αλλοίωσης μέσα συνδυαστικά μέσα από δυο κυρίως παραμέτρους που καθορίζουν το ποσοστό των πλειάδων για αλλοίωση (π%) καθώς και τον αριθμό των LSBs (ξ).

ΑΑ	**Ενέργεια**	**Κωδικός**	**π%**	**γ**	**ξ**
1	Προσθήκη σε 100/γ % πλειάδων	AWM		✓	✓
2	Επίθεση BIT0 π% πλειάδων	BIT0	✓		✓
3	Επίθεση BIT1 π% πλειάδων	BIT1	✓		✓
4	Επίθεση BITFLIP π% πλειάδων	BITFLIP	✓		✓
5	Επίθεση BITRAND π% πλειάδων	BITRAND	✓		✓
6	Στρογγυλοποίηση	RND	✓		✓
7	Μετασχηματισμός Y=γ*X+ξ	TRAN		✓	✓
8	Διαγραφή υποσυνόλου π% πλειάδων	DEL	✓		
9	Ψεύτικη προσθήκη υδατοσήματος σε 100/γ % πλειάδων	CFO		✓	✓
10	Αντιμετάθεση π% πλειάδων	SHF	✓		

Πίνακας 16: Η χρήση των παραμέτρων στις επιθέσεις

Η ρύθμιση των παραπάνω παραμέτρων [21] επηρεάζουν την ανθεκτικότητα, το αριθμό των σφαλμάτων και την αποτελεσματικότητα των ανιχνεύσεων.

Μεταβολή	**Επηρεάζει**
γ↘	ανθεκτικότητα↗ παραμόρφωση↗
ν↗	ανθεκτικότητα↗
ξ↗	ανθεκτικότητα↗ παραμόρφωση↗

Πίνακας 17: Η επιρροή των παραμέτρων στα υδατοσήματα

Τα σενάρια που σχεδιάστηκαν και εκτελέσθηκαν με συνδυασμούς παραμέτρων ξ, γ, και π% φαίνονται στον πίνακα 18 ενώ όλα τα scripts δημιουργίας και εκτέλεσης είναι στο παράρτημα Ι.

Ομάδα	**ΣΕΝΑΡΙΑ (Ακολουθίες ενεργειών)**	**Συνδυασμοί Μεταβλητών**
A	1.Add WM -> Subset Deletion Attack -> Detect WM 2.Add WM -> Bit Flip Attack -> Detect WM 3.Add WM -> Bit Rnd Attack -> Detect WM 4.Add WM -> Shuffle Attack -> Detect WM	**Σταθερές : ξ=5, ν=4** **Συνδυασμοί Μεταβλητών** **γ=50,20,10,8,5,4,2** **%επίθεσης=10,20, … , 90**
B	1. Add WM -> Subset Deletion Attack -> Detect WM 2.Add WM -> Bit Flip Attack -> Detect WM 3.Add WM -> Bit Rnd Attack -> Detect WM 4.Add WM -> Shuffle Attack -> Detect WM	**Σταθερές : %επίθεσης=20%, ξ2=6** **Συνδυασμοί Μεταβλητών** **γ=50,20,10,8,5,3** **ξ=4,5,6,7,8,9**
C	1. Add WM -> Subset Deletion Attack -> Detect WM 2.Add WM -> Bit Flip Attack -> Detect WM 3.Add WM -> Bit Rnd Attack -> Detect WM 4.Add WM -> Shuffle Attack -> Detect WM	**Σταθερές : γ=5, ν=4** **Συνδυασμοί Μεταβλητών** **ξ=3,4,5,6,7,8** **%επίθεσης=10,20, … , 90**
D	Add WM -> CFO Attack -> Detect WM	**Σταθερές : ξ=6, ν=4** **γ=50,20,10,8,5,4,2** **ξ2=2,3,4,5** **γ2=5,10,15,20,25,30,35,40,45**
E	Add WM -> Subset Deletion Attack -> Detect WM Add WM -> Bit 0 Attack -> Detect WM Add WM -> Bit 1 Attack -> Detect WM Add WM -> Bit Flip Attack -> Detect WM Add WM -> Bit Rand Attack -> Detect WM Add WM -> Rounding -> Detect WM Add WM -> Transformation -> Detect WM Add WM -> Shuffle Attack -> Detect WM Add WM -> CFO Attack -> Detect WM (ξ2=4)	**Σταθερές : ξ=5** **Συνδυασμοί Μεταβλητών** **γ=20,10,5** **%επίθεσης=10,20, … , 90**
F	Add WM -> Deletion -> Bit Flip (ξ2=7) -> Shuffle -> Detect WM	**Σταθερές : ξ=6, ν=4** **Συνδυασμοί Μεταβλητών**

Ομάδα	ΣΕΝΑΡΙΑ (Ακολουθίες ενεργειών)	Συνδυασμοί Μεταβλητών
	(Συνδυασμός πολλαπλών επιθέσεων)	**γ=50,20,10,8,5,4,2** **%επίθεσης=10,20, … , 90**
G	Add WM -> CFO Attack -> Detect WM	**Σταθερές : γ=20, ν=4** **Συνδυασμοί Μεταβλητών** **ξ=3,4,5,6,7,8,9,10** **γ2=8,16,24,32** **ξ2=2,3,4,5,6,7,8,9,10**
H	Add WM **weight/attribute** -> BitFlipAttack -> Detect WM	**Σταθερές : ξ=8, ν=4** **Βαρύτητες πιθανότητας επιλογής χαρακτηριστικών [10,6,3,1]** **Συνδυασμοί Μεταβλητών** **γ=50,20,10,8,5,4,2** **%επίθεσης=10,20, … , 90**
I	Add WM **ξ/attribute** -> BitFlipAttack -> Detect WM	**Σταθερές : ν=4** **ξ/χαρακτηριστικό [8,4,2,1]** **Συνδυασμοί Μεταβλητών** **γ=50,20,10,8,5,4,2** **%επίθεσης=10,20, … , 90**
J	Brute Force Watermark creation key guessing	**Σταθερά: creation key=gHUI!216** **Δοκιμές με** **detection key= gHUI!21{N}** **{N}=1,2 … 20**

Πίνακας 18: Τα σενάρια επιθέσεων που υλοποιήθηκαν

Η εκτέλεση των παραπάνω σεναρίων απαίτησε ένα μεγάλο αριθμό υδατογραφήσεων, επιθέσεων και ανιχνεύσεων με συνδυασμούς ξ, γ και ποσοστών αλλοίωσης, όπως φαίνεται στον πίνακα 19. Συνολικά για τα τρία datasets το παραμετρικό πρόγραμμα εκτελέστηκε 3x3654=**10,962 φορές** και δημιουργήθηκαν 3x30=90 γραφήματα από τα οποία βγήκαν χρήσιμα συμπεράσματα.

ΑΑ	Ομάδα Σεναρίων	Υδατογραφήσεις	Επιθέσεις	Ανιχνεύσεις	Σύνολα
1	**A**	7	252	252	**511**
2	**B**	36	144	144	**324**
3	**C**	6	216	216	**438**
4	**D**	6	252	252	**510**
5	**E**	3	243	243	**489**
6	**F**	7	189	189	**385**
7	**G**	8	288	288	**584**

ΑΑ	Ομάδα Σεναρίων	Υδατογραφήσεις	Επιθέσεις	Ανιχνεύσεις	Σύνολα
8	**H**	7	63	63	**133**
9	**I**	7	63	63	**133**
10	**J**	7	-	140	**147**
	Σύνολα	**94**	**1710**	**1850**	**3654**

Πίνακας 19: Αριθμητικά στοιχεία υδατογραφήσεων - επιθέσεων - ανιχνεύσεων σε κάθε dataset

6.4 Η επιρροή των παραμέτρων στα πειράματα

Οι επόμενες υποενότητες σχολιάζουν την επίδραση των γ, ν ,ξ, του μεγέθους του dataset και του τύπου της επίθεσης στην ανθεκτικότητα του υδατοσήματος όπως προκύπτει από τη βιβλιογραφία, τα πειράματα και τα γραφήματα της εργασίας. Ο αλγόριθμος υδατογράφησης της υλοποίησης [21] δεν επηρεάζεται από διαγραφές, ταξινομήσεις ή αντιμεταθέσεις πλειάδων αφού στη δημιουργία των τυχαίων αριθμών (seed) συμμετέχει το κλειδί της κάθε πλειάδας.

6.4.1 Το μέγεθος του υδατοσήματος

Το μέγεθος ενός υδατοσήματος καθορίζεται από την παράμετρο γ, αφού σε κάθε επιλεγμένη πλειάδα επηρεάζεται 1 bit ενός τυχαίου χαρακτηριστικού και ορίζεται ως : **Μ = πλειάδες/γ**. Για παράδειγμα με γ=5 επηρεάζονται 20% των πλειάδων. Σε σύνολο 100,000 πλειάδων τα bits που επηρεάζονται είναι : 100.000/5 = 20.000~ 19Kbits ~ 2,4Kbytes. Αν κάποιος αλγόριθμος επηρεάσει ν ιδιότητες ταυτόχρονα ανά πλειάδα, το μέγεθος Μ θα γίνει Μ*ν.

6.4.2 Η πυκνότητα του υδατοσήματος (γ, ν) και των επιθέσεων

Εκτελέστηκαν πολλά σενάρια επιθέσεων όλων των τύπων, αλλάζοντας μεμονωμένα BITS για συνδυασμούς γ και ποσοστών επιθέσεων διατηρώντας σταθερό το ξ. Οι επιθέσεις bitflip φάνηκαν να είναι οι πιο επιζήμιες αλλά τα αραιά υδατογραφήματα (με μεγάλο ν και γ) είναι ανθεκτικότερα.

6.4.3 Ο αριθμός των LSB (ξ)

Εκτελέστηκαν επίσης, πολλά σενάρια επιθέσεων μεταβολής bits για συνδυασμούς ξ δημιουργίας, ποσοστών και ξ επιθέσεων (ξ2), διατηρώντας σταθερό το γ. Φάνηκε ότι όσο αυξάνεται το ξ (αριθμός LSB) μεγαλώνει η αλλοίωση των τιμών σε βαθμό **$(-2^{\xi}+1 \ldots 2^{\xi}-1)$**. Επίσης, η ανθεκτικότητα της υδατοσήμανσης είναι ανάλογη του ξ αφού τα περισσότερα LSBs μειώνουν την πιθανότητα επιρροής bits από επίθεση.

6.4.4 Η σχέση των αλλοιωμένων πλειάδων με το μέγεθος του υδατογραφήματος

Το μέγεθος ενός υδατογραφήματος καθορίζεται κυρίως από την παράμετρο γ ενώ το μέγεθος της παραμόρφωσης των τιμών μιας ιδιότητας καθορίζεται από το ξ. Συνδυάζοντας αυτά τα δύο, η συνολική αλλοίωση των δεδομένων μιας επίθεσης εξαρτάται από τον αριθμό των πλειάδων που επιλέγονται, από το ξ και από τα επιλεγμένα χαρακτηριστικά που οι τιμές τους διαφέρουν σε τάξη μεγέθους. Η συνολική παραμόρφωση είναι **Π = πλειάδες*ξ/γ**. Άν οι τιμές των αριθμητικών πεδίων είναι 7ψήφοι, τα δεδομένα αλλοιώνονται κατά ένα 1-2% με 5-7 LSBs. Σε 2ψήφιους

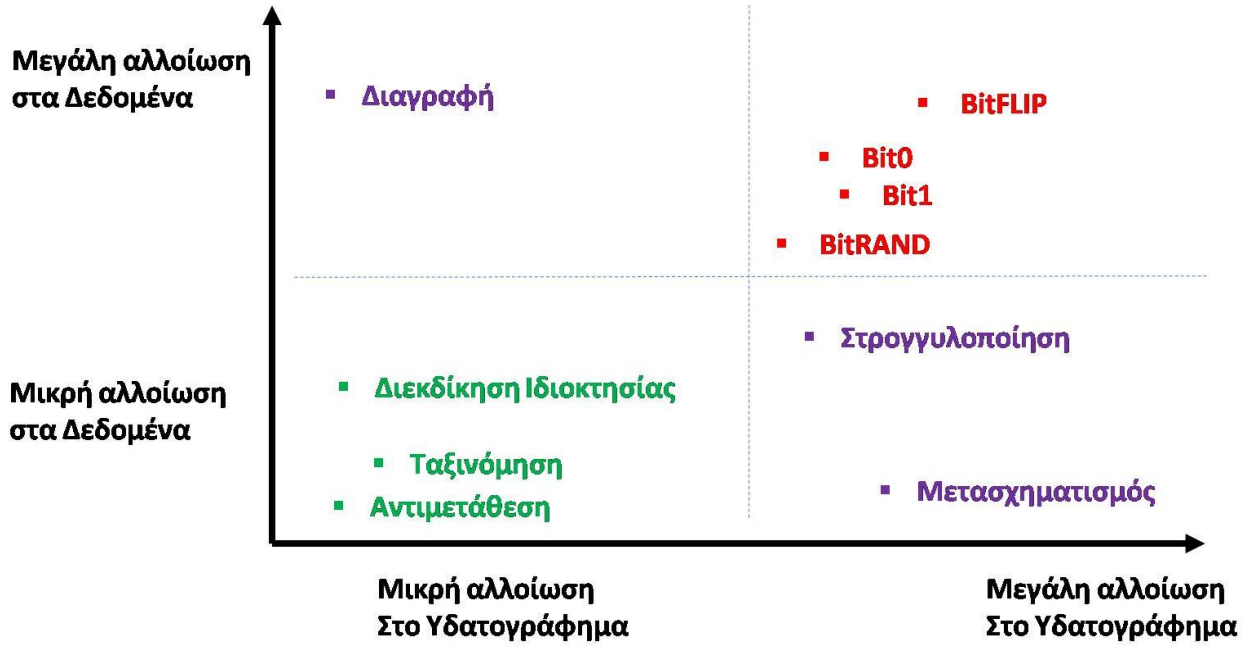
αριθμούς υπάρχει περιορισμός στα 2-3 LSBs. Πρέπει να γίνεται προεπισκόπηση των δεδομένων για την ορθή επιλογή του ξ. **Στην παραλλαγή του αλγόριθμου που επιλέγεται διαφορετικό ξ σε κάθε χαρακτηριστικό, η αλλοίωση των δεδομένων είναι καλύτερα ελεγχόμενη**. Επίσης, βγαίνει ένα χρήσιμο συμπέρασμα ότι από το υποσύνολο μιας βάσης δεδομένων μπορεί να εκτιμηθεί η ανθεκτικότητα για το σύνολο της βάσης.

6.4.5 Ο τύπος της επίθεσης

Μέσα από τις επιθέσεις που πραγματοποιήθηκαν με διαφοροποίηση των παραμέτρων γ, ξ και ν βγαίνει το συμπέρασμα ότι **τη μεγαλύτερη επίπτωση στην αλλοίωση του υδατογραφήματος έχει η επίθεση BITFLIP**, ακολουθούν οι BIT0, BIT1, BITRAND ενώ οι **DEL, SHF και SFO έχουν την ελάχιστη επίπτωση στο υδατογράφημα με βάση τον επιλεγμένο αλγόριθμο [21]**.

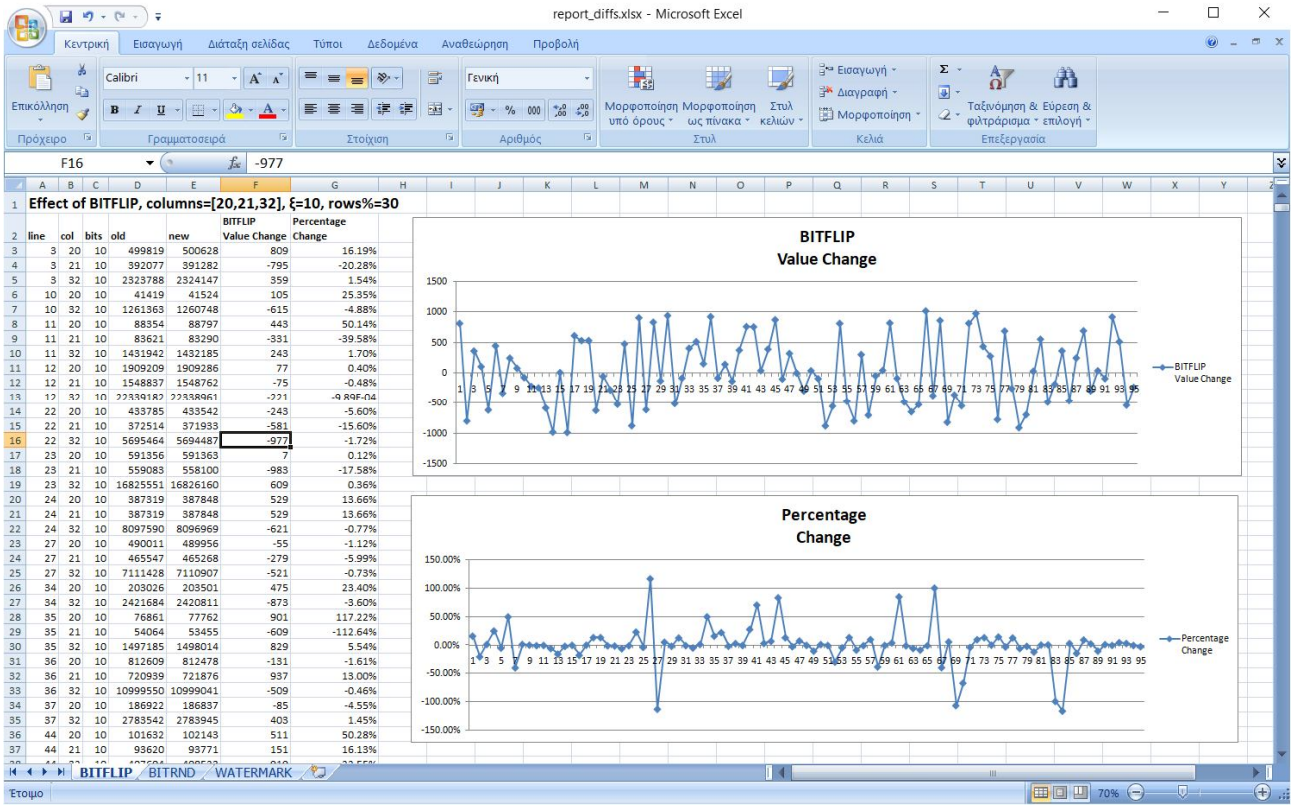
6.4.6 Σύγκριση των επιθέσεων

Με βάση τις προαναφερόμενες παραμέτρους, οι επιθέσεις στον αλγόριθμο υδατογράφησης [21] της εργασίας τοποθετούνται προσεγγιστικά σε διάγραμμα που συγκρίνει το βαθμό αλλοίωσης των δεδομένων σε σχέση με το βαθμό αλλοίωσης του υδατογραφήματος. Προφανώς το μέγεθος της αλλοίωσης καθορίζεται κυρίως από το είδος της επίθεσης και από τον αριθμό των LSBs [ξ].

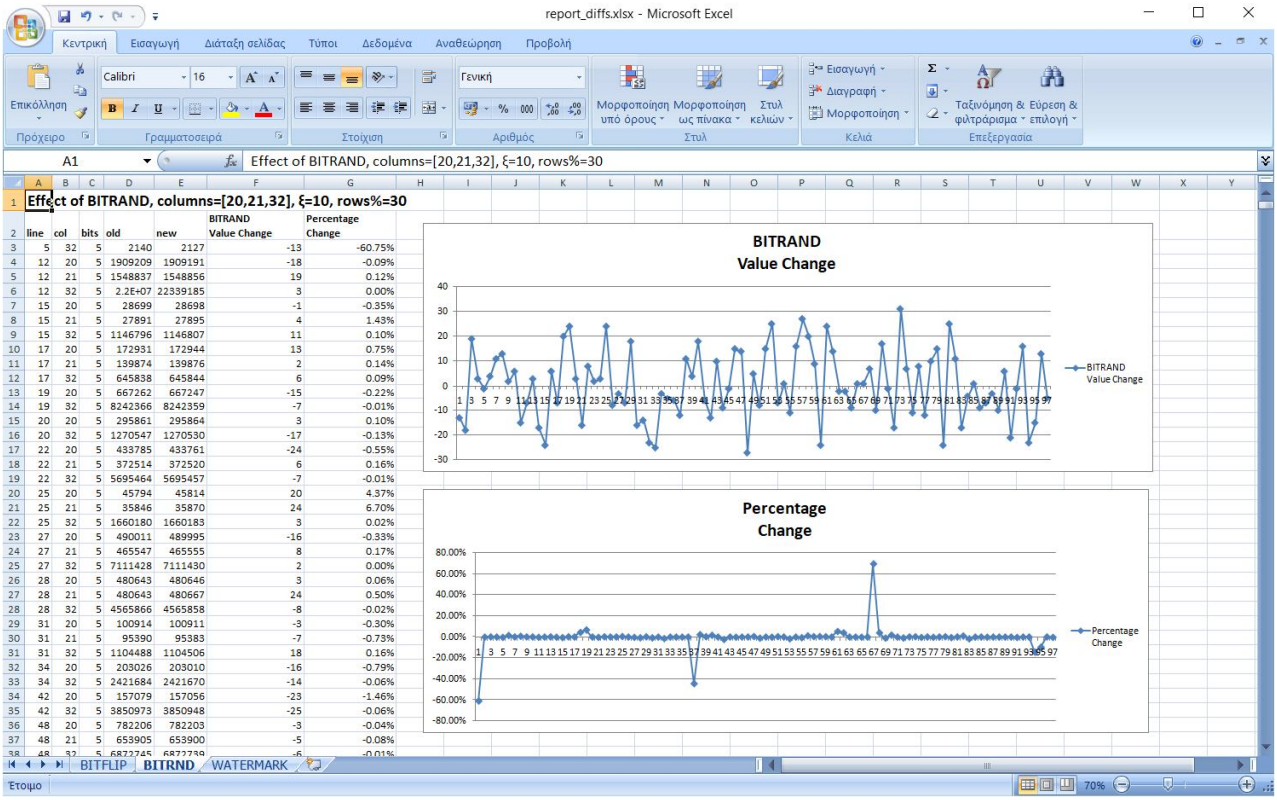


Εικόνα 26: Αλλοίωση στα δεδομένα vs αλλοίωση στο υδατογράφημα

Τα επόμενα γραφήματα, συγκρίνουν την έκταση της αλλοίωσης που επιφέρουν οι δυο επιλεγμένες επιθέσεις BITFLIP και BITRAND με τις ίδιες παραμέτρους (π%, ν, ξ) αναδεικνύοντας τη μεγαλύτερη επίδραση της μορφής BITFLIP.

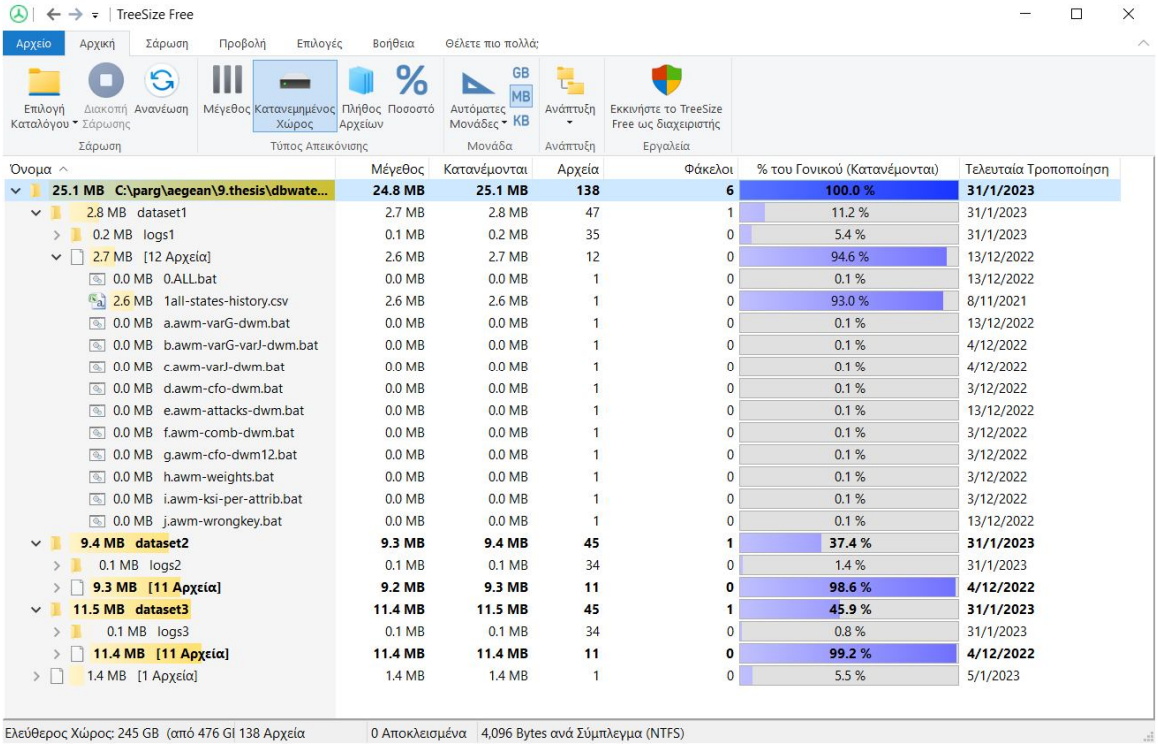


Εικόνα 27: Αλλοίωση δεδομένων από επίθεση BITFLIP

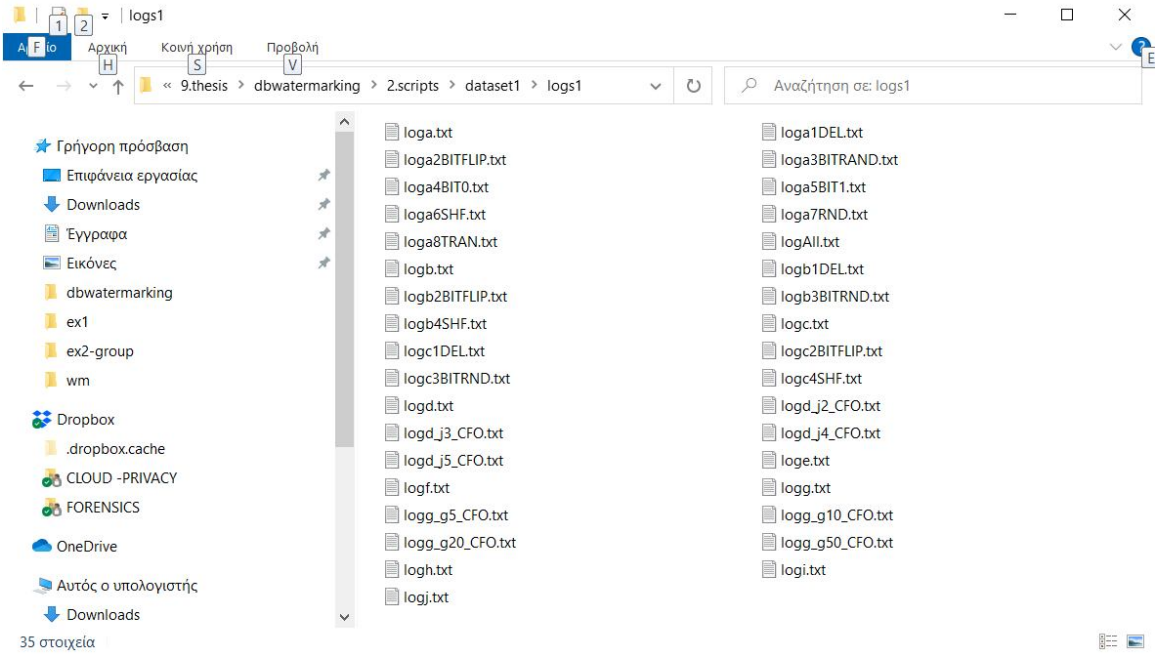


Εικόνα 28: Αλλοίωση δεδομένων από επίθεση BITRAND

Τα πειράματα που έγιναν στα τρία datasets οργανώθηκαν σε φακέλους (Εικ.29), στη συνέχεια παραμετροποιήθηκαν και εκτελέστηκαν τα scripts παράγοντας αρχεία καταγραφής (Εικ.30).



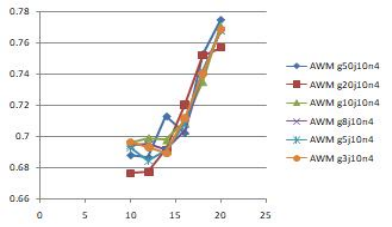
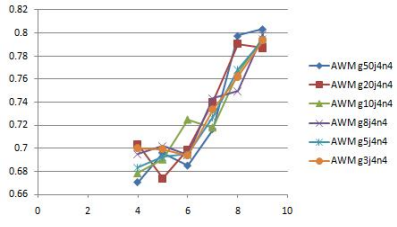
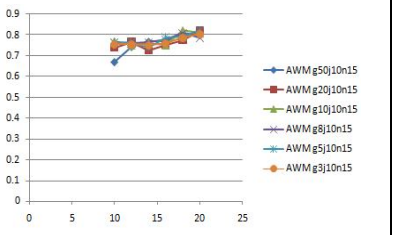
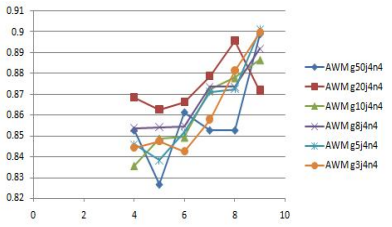
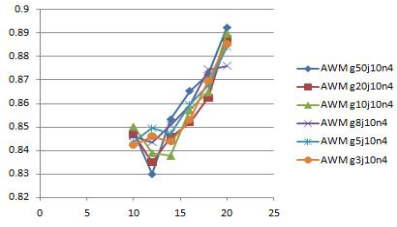
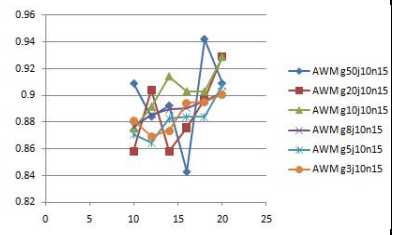
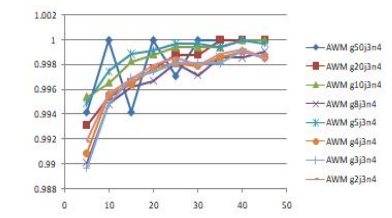
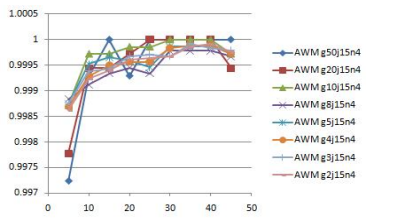
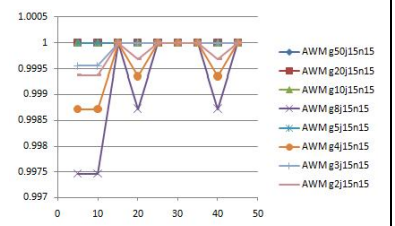
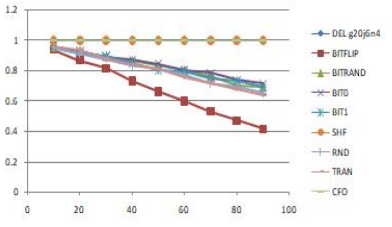
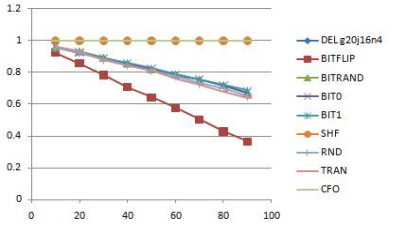
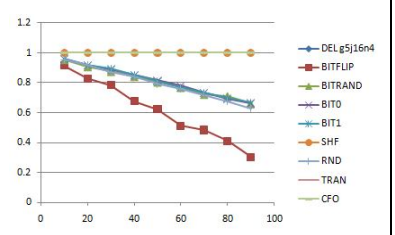
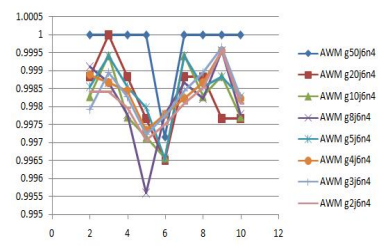
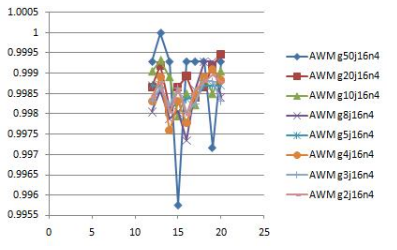
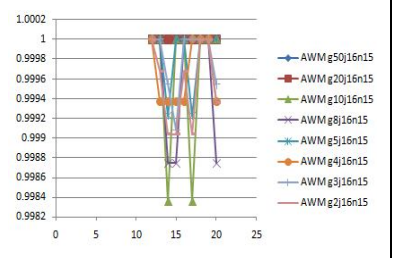
Εικόνα 29: Οργάνωση των σεναρίων και των πειραμάτων



Εικόνα 30: Παραγωγή και ονοματοδοσία των αρχείων καταγραφής - logs

Τα scripts και 90 γραφήματα όλων των πειραμάτων, περιλαμβάνονται στο Παράρτημα Ι.

Τα επόμενα δείγματα γραφημάτων (Πίνακας 20) αποτυπώνουν την επιτυχία της ανίχνευσης στον ψ άξονα και διαφορετικά χρώματα για κάθε συνδυασμό παραμέτρων δημιουργίας της υδατοσήμανσης. Για παράδειγμα η περιγραφή «AWM g50j10n4» της ΜΠΛΕ γραμμής, σημαίνει την αρχική δημιουργία του watermark (Add Watermark) με παραμέτρους γ=50, ξ=10, ν=4.

Dataset **Scenario**	**covid-in-usa-states**	**cryptocurrency-historical-prices-dataset**	**worldwide-stock-market-indices-data**
B2 **BITFLIP** **Attack**	 χ άξονας : ξ2 επίθεσης	 χ άξονας : ξ2 επίθεσης	 χ άξονας : ξ2 επίθεσης
B3 **BITRND** **Attack**	 χ άξονας : ξ2 επίθεσης	 χ άξονας : ξ2 επίθεσης	 χ άξονας : ξ2 επίθεσης
D **CFO** **Attack**	 ξ2=3, χ άξονας : γ2 επίθεσης	 ξ2=16, χ άξονας : γ2 επίθεσης	 ξ2=12, χ άξονας : γ2 επίθεσης
E **ALL** **Attack** **Types**	 ξ2=5, χ άξονας : ποσοστό επίθεσης	 ξ2=15, χ άξονας : ποσοστό επίθεσης	 ξ2=15, χ άξονας : ποσοστό επίθεσης
G **CFO** **Attack**	 γ2=10, χ άξονας : ξ2 επίθεσης	 γ2=5, χ άξονας : ξ2 επίθεσης	 γ2=5, χ άξονας : ξ2 επίθεσης

Πίνακας 20: Δείγματα γραφημάτων από την εκτέλεση σεναρίων στα datasets

7 Συμπεράσματα

Οι καινοτομίες στην εκπόνηση της εργασίας αφορούν στην υλοποίηση ενός **παραμετρικού λογισμικού υδατογράφησης**, επιθέσεων και ανιχνεύσεων που επεξεργάζεται σύνολα δεδομένων (CSV) **χωρίς την απαίτηση βάσης δεδομένων** και στην **ανάπτυξη και εκτέλεση πολλαπλών σεναρίων επιθέσεων με αυτοματοποιημένο τρόπο** μέσα από εκτελέσιμες δέσμες ενεργειών με διαφοροποιημένες τιμές παραμέτρων σε τρία σύνολα δεδομένων αριθμητικών τιμών με **ακέραιους και δεκαδικούς** αριθμούς.

7.1 Οι απαντήσεις στα ερευνητικά ερωτήματα

Ο επόμενος πίνακας συνοψίζει τις απαντήσεις στα ερευνητικά ερωτήματα της εργασίας.

ΑΑ	Ερώτηση	Απάντηση
1	Υπάρχουν νομικά και τεχνολογικά μέτρα προστασίας της πνευματικής ιδιοκτησίας συλλογών δεδομένων ;	Η ενότητα 3.1 αναφέρει τη νομοθεσία που διέπει τα πνευματικά δικαιώματα. Σε επίπεδο Ευρωπαϊκής ένωσης, έχουν εκδοθεί έξι οδηγίες, υπάρχουν αντίστοιχα Ελληνικοί νόμοι, ενώ το Σύνταγμα και ο Χάρτης Θεμελιωδών Δικαιωμάτων περιλαμβάνουν σχετικά άρθρα. Τα τεχνολογικά μέτρα επικεντρώνονται στο υδατογράφημα που περιγράφεται στην ενότητα 3.3 και στα συστήματα διαχείρισης δικαιωμάτων που περιγράφονται στην ενότητα 3.4. Ο συνδυασμός των νομικών και τεχνολογικών μέτρων επιφέρει καλύτερα αποτελέσματα αφού εμποδίζει την πειρατεία με τη χρήση τεχνολογικών εργαλείων και ταυτόχρονα την αποθαρρύνει υπό τον φόβο των ποινών.
2	Υπάρχει τρόπος αναγνώρισης της παράνομης διακίνησης αντιγράφων συλλογών;	Η υδατογράφηση αν χρησιμοποιηθεί ως αποτύπωμα σε κάθε αντίγραφο που διακινείται, μπορεί να βοηθήσει στην ιχνηλάτιση του πειρατή και να αποτελέσει τα απαραίτητα αποδεικτικά στοιχεία ενοχοποίησής του. Μέρος της βιβλιογραφίας [23, 53, 66, 70, 71] αφορά τα αποτυπώματα με προτεινόμενες αρχιτεκτονικές που διακινούν αντίγραφα σε πελάτες με ελεγχόμενο τρόπο προσθέτοντας αόρατα αποτυπώματα.
3	Ποιες είναι οι	Η απάντηση σε αυτό το ερώτημα δεν δίνεται σε απόλυτα μεγέθη (των γ, ξ,

ΑΑ	Ερώτηση	Απάντηση
	κατάλληλες παράμετροι για τη δημιουργία ενός υδατογραφήματος ανθεκτικού σε διάφορες μορφές επιθέσεων ;	ν) στη βιβλιογραφία ούτε προέκυψε από τα πειράματα της εργασίας, αφού εξαρτάται από τα δεδομένα, τον αλγόριθμο υδατογράφησης και την αλλοίωση της επίθεσης. Τα συμπεράσματα των πειραμάτων συνοψίζονται στα παρακάτω σημεία: - Όσο μεγαλύτερη είναι η βάση δεδομένων (σε πλειάδες, αριθμητικά χαρακτηριστικά και απόλυτες τιμές μέσα σε αυτά), τόσο ευκολότερη είναι η διασπορά του υδατοσήματος και η προστασία της. Η υδατογράφηση θα πρέπει να λαμβάνει υπόψη τα χαρακτηριστικά του συνόλου δεδομένων για να έχει μεγαλύτερη ανθεκτικότητα. - Ούτε ο ιδιοκτήτης αλλά ούτε και ο επιτιθέμενος επιθυμούν τη μεγάλη αλλοίωση των δεδομένων. Αν ο επιτιθέμενος υπερεκτιμήσει το ξ, εισάγει μεγάλο σφάλμα χωρίς να αυξάνει την επιτυχία της επίθεσης, επίσης αν υποεκτιμήσει το ξ η επιτυχία θα είναι μικρή - Μια επίθεση αναστροφής bits είναι ισχυρότερη από την τυχαία ανάθεση bits. Μια επίθεση bits που αλλοιώνει ξ LSBs σε μια ιδιότητα ανά πλειάδα, επιφέρει αλλοίωση δεδομένων [A=M*ξ/γ], δηλαδή ξ φορές το μέγεθος του αρχικού υδατογραφήματος [M/γ]. - Αλγόριθμοι όπως ο [21] που βασίζουν τους τυχαίους αριθμούς στο βασικό κλειδί, δεν επηρεάζονται από επιθέσεις διαγραφής, ταξινόμησης η αντιμετάθεσης των στοιχείων. - Η χρήση περισσότερων χαρακτηριστικών (αύξηση ν) στο υδατόσημο δεν αλλάζει το μέγεθος της αλλοίωσης αλλά διαχέει το υδατόσημο καλύτερα και αναγκάζει τον πειρατή να αλλοιώσει μεγαλύτερο μέρος της βάσης, δηλαδή περισσότερα χαρακτηριστικά. - Υπάρχει μια σύνδεση ανάμεσα στο ξ και στο γ. Αν αυξηθεί το γ, μπορεί να αυξηθεί και το ξ κρατώντας σταθερή την αλλοίωση (A=πλειάδες*ξ/γ). Μεγαλύτερα ξ και γ αυξάνουν την ανθεκτικότητα. Συνοψίζοντας, απαιτείται μια αρχική επισκόπηση των δεδομένων, αύξηση των υποψήφιων αριθμητικών χαρακτηριστικά, αύξηση των γ και ξ παράλληλα. Δεδομένα με μικρές αριθμητικές τιμές περιορίζουν τις επιλογές στην τιμή του ξ, ενώ δεδομένα με λίγες πλειάδες περιορίζουν το γ.

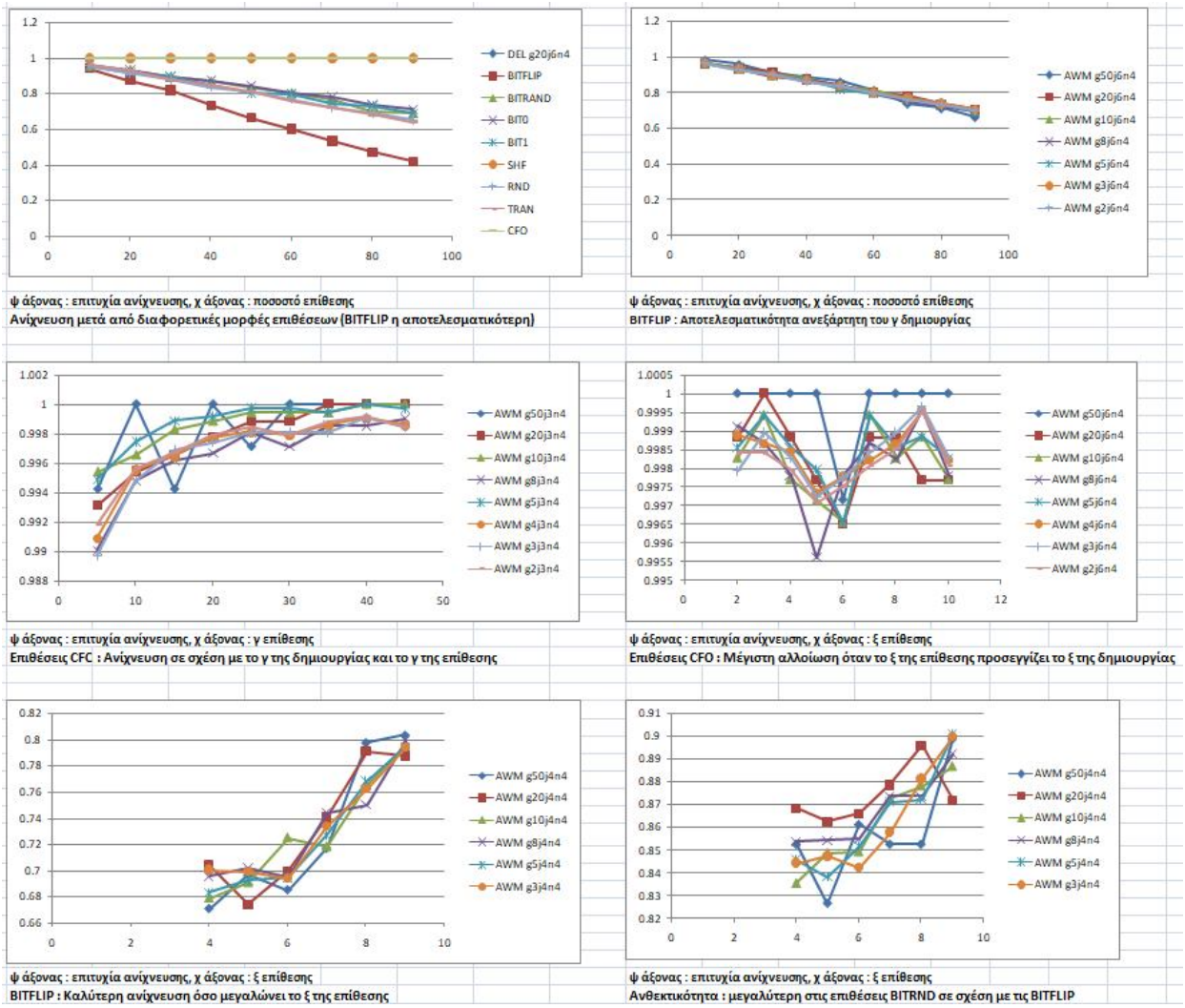
Πίνακας 21: Απαντήσεις στα τρία ερευνητικά ερωτήματα

Ένας επιτιθέμενος ξεκινώντας από μηδενική βάση, δεν μπορεί να γνωρίζει για την ύπαρξη υδατοσήμανσης. Μπορεί να κάνει υποθέσεις για τον αλγόριθμο, υποθέσεις για τα χαρακτηριστικά που επιλέχθηκαν με βάση τα μεγέθη, υποθέσεις για το ξ, υποθέσεις για το γ με βάση την αναλογία ξ/γ, όπου μεγαλύτερο ξ ευνοεί μεγαλύτερο γ, άρα σε μια υποθετική επίθεση η επικέντρωση θα ήταν σε όσο μεγαλύτερο ξ και μικρότερο γ. Αν ο σκοπός της επίθεσης είναι η εξαφάνιση του υδατογραφήματος και ο αλγόριθμος επηρεάζεται από την αρχική ταξινόμηση, επιθέσεις διαγραφής η ανακατέματος πλειάδων είναι μια καλή επιλογή. Αν ο αλγόριθμος είναι ανεξάρτητος της ταξινόμησης [21] οι επιθέσεις BITFLIP και στρογγυλοποίησης επιφέρουν μεγαλύτερη

αλλοίωση. Αν ο στόχος δεν είναι η αλλοίωση του υδατογραφήματος αλλά ο ισχυρισμός της ιδιοκτησίας θα πρέπει η επιλεγμένη μέθοδος επίθεσης να προσθέτει ισχυρό ψεύτικο υδατογράφημα και παράλληλα να αλλοιώνει το αληθινό υδατογράφημα σε ένα βαθμό.

7.2 Τα σημαντικότερα διαγράμματα και συμπεράσματα

Τα σημαντικά διαγράμματα που υποστηρίζουν την απάντηση στο τρίτο ερευνητικό ερώτημα της εργασίας έχουν συγκεντρωθεί εδώ μαζί με κάποια σημαντικά συμπεράσματα. Τα γραφήματα αποτυπώνουν την επιτυχία της ανίχνευσης στον ψ άξονα με διαφορετικά χρώματα για κάθε συνδυασμό παραμέτρων δημιουργίας της υδατοσήμανσης. Για παράδειγμα «AWMg50j3n4» στο τρίτο γράφημα σημαίνει δημιουργία του watermark με παραμέτρους γ=50, ξ=3, ν=4.



Εικόνα 31: Επιλεγμένα γραφήματα στη διεξαγωγή συμπερασμάτων

Συνοψίζονται τα συμπεράσματα που προκύπτουν από τα πειράματα στα τρία σύνολα δεδομένων.

i. Ο αλγόριθμος [21] είναι ανθεκτικός σε επιθέσεις διαγραφής/αντιμετάθεσης πλειάδων.

ii. Αν το κλειδί της ανίχνευσης είναι το ίδιο με το κλειδί της δημιουργίας η επιτυχία στην ανίχνευση είναι 100% διαφορετικά κυμαίνεται γύρω στο 50% , με βάση την πιθανότητα το next(G) του αλγόριθμου να είναι άρτιος ή περιττός αριθμός.

iii. Οι επιθέσεις σε LSBs έχουν άμεσο αντίκτυπο ανάλογο του ξ_2 επίθεσης, ανεξάρτητα του γ.

iv. Η επίθεση CFO αλλοιώνει το αρχικό υδατογράφημα αντιστρόφως ανάλογα με το γ_2 της επίθεσης και μεγιστοποιείται όσο το ξ_2 προσεγγίζει το αρχικό ξ.

v. Οι αποτελεσματικότερες επιθέσεις είναι κατά σειρά οι BITFLIP, BITRND, BIT0, BIT1.

vi. Τα αραιά υδατογραφήματα (με μεγάλο ν και γ) είναι πιο ανθεκτικά στις επιθέσεις.

vii. Η ανθεκτικότητα είναι ανάλογη του ξ και μειώνει την επιρροή επιθέσεων με μικρότερα ξ_2.

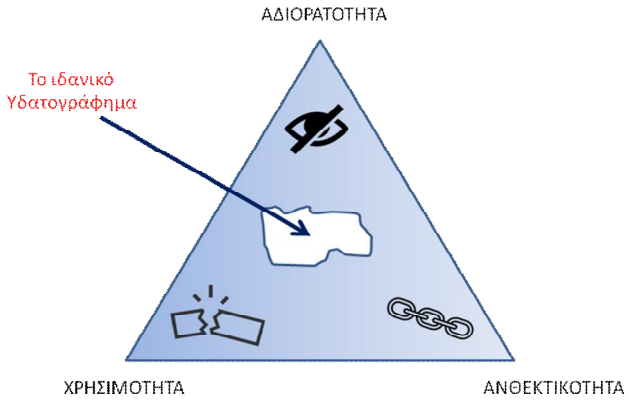
viii. Από ένα υποσύνολο των πλειάδων μπορεί να εκτιμηθεί η ανθεκτικότητα για το σύνολο .των δεδομένων, αυτό ισχύει σε αλγόριθμους ανθεκτικούς σε διαγραφές υποσυνόλου [21].

7.3 Το υδατογράφημα ως μέσο αναγνώρισης του δημιουργού

Όπως αναφέρθηκε στο κεφάλαιο 2, η ισχυρή υδατογράφηση πρέπει να είναι αρκετή ώστε να μην αλλοιώνεται εύκολα αλλά έως ένα όριο που δεν θα αλλοιώνει τα ίδια τα δεδομένα. Αυτό είναι μια ποιοτική προσέγγιση που δεν οδηγεί απαραίτητα στον καθορισμό των ιδανικών παραμέτρων σε μια υδατογράφηση. Από την πλευρά του επιτιθέμενου ο στόχος είναι παρόμοιος, η αλλοίωση του υδατογραφήματος χωρίς την αλλοίωση των δεδομένων. Η χρησιμότητα των δεδομένων και ο βαθμός αλλοίωσης στα αριθμητικά δεδομένα σχετίζεται με μια σειρά παραμέτρους :

- Τα απόλυτα και τα στατιστικά μεγέθη των αριθμητικών δεδομένων
- Το ποσοστό αλλοίωσης που είναι αποδεκτό και σε επίπεδο χαρακτηριστικού
- Τη σχετική χρησιμότητα ανάμεσα στα χαρακτηριστικά

Στη συμμετρική κρυπτογραφία και στις συναρτήσεις κατακερματισμού, η εικόνα είναι ξεκάθαρη, όσο μεγαλώνει το μήκος του κλειδιού ή του hash, έχουμε μεγαλύτερη ασφάλεια με αυξημένο υπολογιστικό κόστος και κόστος αποθήκευσης. Στην περίπτωση της υδατογράφησης, το κατάλληλο σημείο ισορροπίας βρίσκεται ανάμεσα στην ανθεκτικότητα [την ασφάλεια], τη χρησιμότητα [το μέγεθος της αλλοίωσης] και την αδιορατότητα [το επίπεδο της στεγανογραφίας] όπως δείχνει η Εικόνα 32.



Εικόνα 32: Η ισορροπία ανάμεσα στην Ανθεκτικότητα, τη Χρησιμότητα και την Αδιορατότητα

7.4 Διαπιστώσεις, βελτιώσεις και προτάσεις

Τα γενικά συμπεράσματα από τη μελέτη της βιβλιογραφίας και τον πειραματισμό με τα υδατοσήματα, τα σενάρια επιθέσεων και τις ανιχνεύσεις, συνοψίζονται εδώ.

i. Η έρευνα στο χώρο δεν έχει μια κοινή βάση κριτηρίων για την παραμετροποίηση που μπορεί να οδηγήσει στη βέλτιστη ανθεκτικότητα

ii. Κάποιες μορφές επιθέσεων είναι δραστικές και μπορούν να καταστρέψουν τα υδατογραφήματα χωρίς να υπάρχουν γνωστά αντίμετρα

iii. Υπάρχουν παραλλαγές αλγορίθμων που οδηγούν σε πυκνότερα υδατογραφήματα, επηρεάζοντας πολλαπλές στήλες στην ίδια πλειάδα ή πολλαπλά bits στην ίδια ιδιότητα. Η μεγάλη διασπορά του υδατοσήματος σε πλειάδες, χαρακτηριστικά και LSBs φαίνεται να είναι πιο σημαντική από τη μεγάλη πυκνότητα στην ανθεκτικότητα.

iv. Κάποιες μέθοδοι χρησιμοποιούν δυαδικές υπογραφές βιομετρικών στοιχείων, εικόνας ή σήματος όμως ο ιδανικός βαθμός της παραμόρφωσης εξαρτάται από τα δεδομένα και παραμένει ένα ζήτημα για περισσότερη διερεύνηση.

v. Τα δεδομένα που περιλαμβάνουν κείμενο, κατηγορικές και ανωνυμοποιημένες πληροφορίες είναι περισσότερο ευαίσθητα στην προσθήκη θορύβου αφού οι παρεμβάσεις γίνονται ευκολότερα αντιληπτές. Εκεί, μια ταξινόμηση ή μεταβολή δεδομένων μέσα από διαγραφές και εισαγωγές πλειάδων αρκούν για να καταστρέψουν ένα υδατογράφημα.

vi. Οι βιβλιογραφικές πηγές που προτείνουν, υλοποιούν, δοκιμάζουν και συγκρίνουν σχήματα υδατογράφησης αναφέρουν πειράματα και προβάλουν ανάλυση των αποτελεσμάτων αλλά δεν δημοσιεύουν την οργάνωση των πειραμάτων αυτών ώστε να επαναληφθούν εύκολα στο μέλλον. Στον πραγματικό κόσμο οι επιθέσεις μπορεί να συνδυάσουν πολλές τεχνικές ταυτόχρονα με στόχο την αλλοίωση του υδρογραφήματος. Η εργασία αυτή οργάνωσε τα σενάρια μέσα από εκτελέσιμα scripts για να μπορούν να επαναληφθούν και να επεκταθούν εύκολα στο μέλλον ενώ παρείχε γραφική απεικόνιση όλων των πειραμάτων.

vii. Η εργασία ανέδειξε τη μεγάλη χρησιμότητα των εργαλείων υδατογράφησης καθώς και την αξία της οπτικής απεικόνισης διαδικασιών και αποτελεσμάτων.

viii. Η ερευνητική κοινότητα χρειάζεται περισσότερα εργαλεία σύγκρισης τεχνικών και μετρήσεων απόδοσης των σχημάτων υδατογράφησης αλλά και εκπαιδευτικά εργαλεία που να χρησιμοποιούν το μέσο της οπτικοποίησης για τις διάφορες μορφές της υδατογράφησης σε σύνολα δεδομένων.

Η υδατογράφηση με την ενσωμάτωση ενός κρυφού μηνύματος, μπορεί να ανιχνεύσει αλλοιώσεις (στα εύθραυστα σχήματα), να επιβεβαιώσει το δημιουργό (στα ισχυρά σχήματα) ή να ιχνηλατήσει τον πειρατή (με αποτύπωμα) όπως ανέλυσε η εργασία. Ένα πνευματικό έργο είναι άυλο στη φύση του, διαμορφώνεται από την ανθρώπινη δημιουργικότητα την έμπνευση, είναι μοναδικό όπως και ο δημιουργός του, αξίζει το σεβασμό, την αναγνώριση και την απόδοση των δικαιωμάτων που του αναλογούν και πρέπει να προστατεύεται με τεχνολογικά και νομικά μέσα.

Η ασφάλεια, κατέχει τη δεύτερη θέση στην ιεραρχία των ανθρώπινων αναγκών της πυραμίδας Maslow (Εικόνα 33). Η ασφάλεια διαχρονικά είχε πολλές ερμηνείες. Στη σύγχρονη εποχή που ζούμε, η ασφάλεια συνδέεται άρρηκτα με την προστασία της ελευθερίας, το σεβασμό

των ατομικών και των κοινωνικών δικαιωμάτων, τη διαφύλαξη των προσωπικών δεδομένων, την ιδιωτικότητα και την προστασία των πνευματικών δικαιωμάτων.

Εικόνα 33: Η ασφάλεια στη δεύτερη θέση της πυραμίδα των ανθρώπινων αναγκών (Maslow)

Βελτιώσεις στις επόμενες εκδόσεις του παραμετρικού λογισμικού μπορούν να περιλαμβάνουν την υλοποίηση περισσότερων αλγορίθμων υδατογράφησης ΒΔ, μετρήσεις απόδοσης και συγκρίσεις ανθεκτικότητας ανάμεσα σε διαφορετικά σχήματα, αριθμητικά δεδομένα μορφών XML ή JSON.

Γενικότερες προτάσεις που απορρέουν από την εκπόνηση της εργασίας αφορούν :

1. Τη χρήση του παρόντος παραμετρικού λογισμικού ως εκπαιδευτικό εργαλείο στη δημιουργία και εκτέλεση σεναρίων σε περιβάλλοντα MS windows και linux.
2. Τη διάθεση των εργαλείων στην ακαδημαϊκή κοινότητα και τη δημοσίευση του έργου.
3. Τη δημιουργία κεντρικού ηλεκτρονικού μητρώου ανοιχτού λογισμικού υδατοσήμανσης δεδομένων και πειραματικών δεδομένων για την online εκτέλεση πειραμάτων με δυνατότητες κατηγοριοποίησης, αναζήτησης, ελέγχου και δοκιμών των υλοποιημένων αλγορίθμων μέσα από υπηρεσία : Experiment as a Service (EaaS).
4. Την επαλήθευση των πειραμάτων και των αποτελεσμάτων της βιβλιογραφίας.
5. Την αυτοματοποιημένη σύγκριση της ανθεκτικότητας και αποδοτικότητας όλων των γνωστών αλγορίθμων υδατοσήμανσης για ΒΔ, βασισμένη σε κοινά αποδεκτούς δείκτες αποδοτικότητας.
6. Την προσθήκη παραδειγμάτων υδατογράφησης συνόλων δεδομένων στο CrypTool.

Η επιστημονική κοινότητα χρειάζεται δεδομένα, εργαλεία και ηλεκτρονικές υπηρεσίες για να υποστηρίζει την πειραματική διαδικασία της έρευνας και να εξάγει αδιάσειστα, τεκμηριωμένα και επαληθεύσιμα συμπεράσματα.

Βιβλιογραφία

[1] Δημήτριος Τσώλης (2015), «Προστασία και Διαχείριση της Πνευματικής Ιδιοκτησίας Ψηφιακού Περιεχομένου στο Διαδίκτυο και τα Σύγχρονα Δίκτυα», Ελληνικά Ακαδημαϊκά Ηλεκτρονικά Συγγράμματα και Βοηθήματα (www.kallipos.gr) [πρόσβαση 24/9/2022]

[2] ΟΔΗΓΙΑ 96/9/ΕΟΚ ΤΟΥ ΕΥΡΩΠΑΪΚΟΥ ΚΟΙΝΟΒΟΥΛΙΟΥ ΚΑΙ ΤΟΥ ΣΥΜΒΟΥΛΙΟΥ της 11ης Μαρτίου 1996 σχετική με τη νομική προστασία των βάσεων δεδομένων

[3] THE CHARTER OF DIGITAL RIGHTS, A guide for policy-makers, EDRi papers, ed.10 https://edri.org/wp-content/uploads/2014/06/EDRi_DigitalRightsCharter_web.pdf [πρόσβαση 24/9/2022]

[4] Κοντόβας Γεώργιος (2012), Μεταπτυχιακή Διατριβή, Πανεπιστήμιο Πειραιώς, «Η προστασία των πνευματικών δικαιωμάτων των δημιουργών έργων ηλεκτρονικών δεδομένων σύμφωνα με το δίκαιο της πνευματικής ιδιοκτησίας», Σεπ 2012 [πρόσβαση 24/9/2022] https://dione.lib.unipi.gr/xmlui/bitstream/handle/unipi/5293/Kontovas.pdf?sequence=2&isAllowed=y

[5] Δ. ΚΑΛΛΙΝΙΚΟΥ (2001), «Η ΠΝΕΥΜΑΤΙΚΗ ΙΔΙΟΚΤΗΣΙΑ ΣΤΙΣ ΑΡΧΕΣ ΤΟΥ 21ου ΑΙΩΝΑ» https://lekythos.library.ucy.ac.cy/bitstream/handle/10797/10852/ilektr009a.pdf?sequence=1

[6] Δ. ΚΑΛΛΙΝΙΚΟΥ (1995), «ΠΝΕΥΜΑΤΙΚΗ ΙΔΙΟΚΤΗΣΙΑ ΚΑΙ ΣΥΓΓΕΝΙΚΑ ΔΙΚΑΙΩΜΑΤΑ ΣΤΗΝ ΕΛΛΑΔΑ», Νομική Πανεπιστημίου Αθηνών [πρόσβαση 24/9/2022] https://lekythos.library.ucy.ac.cy/bitstream/handle/10797/12251/12-13teuxos023.pdf?sequence=1

[7] Scott J. Schackelford (2016), «Protecting Intellectual Property and Privacy in the Digital Age: The Use of National Cybersecurity», Indiana University

[8] Estelle Derclaye (2007), «Intellectual property rights on information and market power – comparing European and American protection of databases»

[9] Philippe Jougleux, «Ανοιχτή πρόσβαση και πνευματική ιδιοκτησία : μια ευαίσθητη ισορροπία;», European University Cyprus

[10] Πρόδρομος Τσιαβός (2014), «Ανοικτή Πρόσβαση Πνευματικά Δικαιώματα Ανοικτές Άδειες», Εθνικό Κέντρο Τεκμηρίωσης

[11] Μαρία Κανελλοπούλου Μπότη, «Χρήση βάσεων δεδομένων σε βιβλιοθήκες και αρχεία», Τμήμα Αρχειονομίας Βιβλιοθηκονομίας, Ιόνιο Πανεπιστήμιο

[12] Παραδείσας Μανώλης (2012), «Ψηφιακή Διαχείριση Πνευματικών Δικαιωμάτων (Digital Rights Management System)», διπλωματική εργασία, Πανεπιστήμιο Πειραιώς

[13] Διονυσία Καλλίνικου, Θεόδωρος Καρούνος, Μαρίνος Παπαδόπουλος, «Οι ελληνοποιημένες Άδειες Creative Commons», 16ο Πανελλήνιο Συνέδριο Ακαδημαϊκών Βιβλιοθηκών «Ο ανθρώπινος παράγοντας στη διαμόρφωση της σημερινής και της μελλοντικής βιβλιοθήκης»

[14] Tabrez Ahmad†* and Sourav Dan, «Comparative Analysis of Copyright Protection of Databases: The Path to Follow», Journal of Intellectual Property Rights, Vol 17, March 2012, pp 111-121

[15] Konstantinos Loukopoulos (2016), «Πνευματική Κοινοκτημοσύνη (αντί για Πνευματική Ιδιοκτησία;» www.lawpractice.gr, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2777515

[16] Πρόδρομος Τσιαβός (2015), «Νομικό Πλαίσιο για την παροχή και συνεισφορά Δεδομένων», Ημερίδα European Language Resource Coordination

[17] Iurisperitus Kiadó Szeged, (2018), «Introduction to Digital Copyright Law», 3rd Edition

[18] Μπαριτάκη Μαρία (2020), "Η ΣΥΓΚΡΟΥΣΗ ΤΗΣ ΠΝΕΥΜΑΤΙΚΗΣ ΙΔΙΟΚΤΗΣΙΑΣ ΜΕ ΤΑ ΠΡΟΣΩΠΙΚΑ ΔΕΔΟΜΕΝΑ", ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝΣΠΟΥΔΩΝ ΤΗΣ ΣΧΟΛΗΣ ΟΙΚΟΝΟΜΙΚΩΝ ΕΠΙΧΕΙΡΗΜΑΤΙΚΩΝ & ΔΙΕΘΝΩΝ ΣΠΟΥΔΩΝ ΤΟΥ ΠΑΝΕΠΙΣΤΗΜΙΟΥ ΠΕΙΡΑΙΩΣ, «ΔΙΚΑΙΟ ΚΑΙ ΟΙΚΟΝΟΜΙΑ – MASTER IN LAW AND ECONOMICS»

[19] Κυριάκος Παπαγεωργίου, "Πνευματική Ιδιοκτησία και User Generated Content", Διπλωματική Εργασία, Πανεπιστήμιο Αιγαίου, Τμήμα Μηχανικών Πληροφοριακών & Επικοινωνιακών Συστημάτων

[20] Yong ZHANG, Xiamu NIU, Dongning ZHAO, «A Method of Protecting Relational Databases Copyright with Cloud Watermark», World Academy of Science, Engineering and Technology 3 2007

[21] R. Agrawal, and J. Kiernan, "Watermarking Relational Databases". In Proceedings of the 28th International Conference on Very Large Databases (VLDB), pp.155-166 (2002)

[22] T. Tzouramanis, "A Robust Watermarking Scheme for Relational Databases", In Proceedings of the 6th IEEE International Conference on Internet Technology & Secured Transactions (ICITST), pp.783-790 (2011). http://ttzouram.users.uth.gr/TechnicalReports/ieee-icitst-2011.pdf [πρόσβαση 1/10/2022]

[23] Ahmed T.Sadiq, Amjed Abbas Ahmed, Bashr Saadoon Mahdi, "Watermarking and Fingerprinting for Relational Database"

[24] Stavros Kyriakopoulos, TheodorosTzouramakis, University of Aegean, Yannis Manolopoulos, Aristotle University of Thessaloniki (2017) , "The dbMark: A Benchmarking System for Watermarking Methods for Relational Databases," [πρόσβαση 1/10/2022] https://datalab-old.csd.auth.gr/wp-content/uploads/publications/RCIS2017ktm.pdf

[25] BrijeshB.Mehta, UdaiPratap Rao, "A Novel approach as Multi-place Watermarking for Security in Database", National Institute of Technology Surat, Gujarat, INDIA, Int'l Conf. Security and Management | SAM'11

[26] M. Kamran and Muddassar Farooq (2018), "A Comprehensive Survey of Watermarking Relational Databases Research", arXiv:1801.08271v1 [cs.CR] 25 Jan 2018 [πρόσβαση 1/10/2022] https://arxiv.org/pdf/1801.08271.pdf

[27] Mustafa Bilgehan Imamoglu, Mustafa Ulutas, and Guzin Ulutas (2016), "A New Reversible Database Watermarking Approach with Firefly Optimization Algorithm", Hindawi Mathematical Problems in Engineering, Volume 2017, Article ID 1387375, 14 pages https://downloads.hindawi.com/journals/mpe/2017/1387375.pdf [πρόσβαση 1/10/2022]

[28] Raju Halder, Shantanu Pal, Agostino Cortesi (2016), "Watermarking Techniques for Relational Databases: Survey, Classification and Comparison", Journal of Universal Computer Science, vol. 16, no. 21 (2010), 3164-3190 [πρόσβαση 1/10/2022] https://eprints.qut.edu.au/209984/1/jucs_16_21_3164_3190_halder.pdf

[29] Wilfred Ng and Ho-Lam Lau, "Effective Approaches for Watermarking XML Data", Department of Computer Science, The Hong Kong University of Science and Technology, Hong Kong https://link.springer.com/chapter/10.1007/11408079_9

[30] Radu Sion, Mikhail Atallah, Sunil Prabhakar (2003), "RIGHTS PROTECTION FOR RELATIONAL DATA", CERIAS Tech Report 2003-11, Center for Education and Research in Information Assurance

and Security, Purdue University West Lafayette, IN 47907, USA [πρόσβαση 1/10/2022] https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.158.7444&rep=rep1&type=pdf

[31] Yingjiu Li, Vipin Swarup, Sushil Jajodia, "Constructing a Virtual Primary Key for Fingerprinting Relational Data" https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.302.6064 [πρόσβαση 1/10/2022]

[32] Namrata Gursale, Arti Mohanpurkar (2014), "A Robust, Distortion Minimization Fingerprinting Technique for Relational Database", International Journal on Recent and Innovation Trends in Computing and Communication, Volume: 2 Issue: 6 [πρόσβαση 1/10/2022]
https://d1wqtxts1xzle7.cloudfront.net/35482811/A_Robust_Distortion_Minimization_Fingerprinting_Technique_for_Relational_Database-with-cover-page-v2.pdf

[33] Buse Gul Atli Tekgul, N. Asokan (2022), "On the Effectiveness of Dataset Watermarking", IWSPA '22, April 27, 2022, Baltimore, MD, USA [πρόσβαση 1/10/2022] https://dl.acm.org/doi/pdf/10.1145/3510548.3519376

[34] Francesc Seb´e, Josep Domingo-Ferrer, and Agust´ı Solanas, Rovira i Virgili, "Noise-Robust Watermarking for Numerical Datasets", University of Tarragona. Dept. of Computer Engineering and Maths, Av. Pa¨ısos Catalans 26, E-43007 Tarragona, Catalonia [πρόσβαση 1/10/2022] https://link.springer.com/chapter/10.1007/11526018_14

[35] Yang Lu a,b, Zongwei Tang a, Xiuli Chai a,b,*, Mingxu Wang a, Shiping Song, "A hierarchical protection scheme for intellectual property of semi-open source datasets based on double watermarking", Optik - International Journal for Light and Electron Optics 269 (2022) 169931

[36] Saman Iftikhar, M Kamran, Zahid Anwar (2015), "A survey on reversible warermarking techniques for relational databases", Security and Communication Networks, vol 8, no 15, 2015

[37] Harold F.Tipton, Michi Krause, "Information Security Management Book", 6th edition

[38] YUICHI NAKAMURA 1 AND HIROAKI NISHI (2021), "Digital Watermarking for Anonymized Data With Low Information Loss", IEEE Access https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9530662 [πρόσβαση 1/10/2022]

[39] Yuichi NAKAMURA, Yoshimichi NAKATSUKA and Hiroaki NISHI "Novel Method to Watermark Anonymized Data for Data Publishing", Special Section on Information and Communication System Security, IEICE TRANS. INF. & SYST., VOL.E100–D, NO.8 AUGUST 2017 [πρόσβαση 1/10/2022] https://www.jstage.jst.go.jp/article/transinf/E100.D/8/E100.D_2016ICP0015/_pdf

[40] Rajesh Kumar Tiwari & G. Sahoo (2013), "A Novel Watermarking Scheme for Secure Relational Databases", Information Security Journal: A Global Perspective, 22:3, 105-116, DOI: 10.1080/19393555.2013.801538 [πρόσβαση 1/10/2022] https://www.tandfonline.com/doi/pdf/10.1080/19393555.2013.801538?needAccess=true

[41] Sanjay Kumar1 · Binod Kumar Singh1 · Mohit Yadav1 (2020), "A Recent Survey on Multimedia and Database Watermarking", Multimedia Tools and Applications (2020) 79:20149–20197 https://link.springer.com/content/pdf/10.1007/s11042-020-08881-y.pdf [πρόσβαση 1/10/2022]

[42] Jian Zhao (1996), "A WWW SERVICE TO EMBED AND PROVE DIGITAL COPYRIGHT WATERMARKS", Fraunhofer Institute for Computer Graphics Wilhelminenstr. 7, 64283 Darmstadt GERMANY

[43] Farah Naz & Abid Khan & Mansoor Ahmed & Majid Iqbal Khan & Sadia Din & Awais Ahmad & Gwanggil Jeon (2018), "Watermarking as a service (WaaS) with anonymity", Multimedia Tools and Applications (2020) 79:16051–16075

[44] Ali Al-Haj, Ashraf Odeh, and Shadi Masadeh, "Copyright Protection of Relational Database Systems", 2nd International Conference, Networked Digital Technologies 2010, Prague, Czech Republic

[45] Michael Gertz, Sushil Jajodia (2008), "Handbook of database security, Applications and Trends", Springer Radu Sion, "Database Watermarking for Copyright Protection", Yingjiu Li, "Database Watermarking: A Systematic View"

[46] S.A. Shah, Sun Xingming and Hamadou Ali , Majid Abdul (2010), «Query Preserving Relational Database Watermarking», Pervasive Computing Lab Zhejiang University Hangzhou, Zhejiang P.R. China https://www.informatica.si/index.php/informatica/article/download/363/364 [πρόσβαση 1/10/2022]

[47] Vidhi Khanduja (2017), «Database watermarking, a technological protective measure: Perspective, security analysis and future directions», Journal of Information Security and Applications 37 (2017) 38–49, https://www.sciencedirect.com/science/article/pii/S2214212617300443, [πρόσβαση 1/10/2022]

[48] Yexing Zhang1, Zihou Wang4, Zhaoguo Wang2,3(B), and Chuanyi Liu2,3 (2021), «A Robust and Adaptive Watermarking Technique for Relational Database», Cyber Security, 18th China Annual Conference, CNCERT 2021

[49] Udai Pratap Rao a, Dhiren R. Patel a, Punitkumar M. Vikani (2012), «Relational Database watermarking for Ownership protection», 2nd International Conference on Communication, Computing & Security [ICCCS-2012]

[50] Javier Franco Contreras (2015), «Watermarking Services for medical database service security», PhD, HAL open science

[51] Δέσποινα Σταπατόρη "Οι βάσεις δεδομένων και η προστασία τους από το δίκαιο πνευματικής ιδιοκτησίας", Μεταπτυχιακή διατριβή ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ , ΝΟΜΙΚΗ ΣΧΟΛΗ , ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ ΤΟΜΕΑΣ ΕΜΠΟΡΙΚΟΥ ΚΑΙ ΟΙΚΟΝΟΜΙΚΟΥ ΔΙΚΑΙΟΥ

[52] Radu Sion, «Database Watermarking», Network Security and Applied Cryptography Lab Computer Science, Stony Brook University

[53] YAXIAO SONG (2004), «Watermarking of relational databases», NATIONAL UNIVERSITY OF SINGAPORE

[54] Σύνταγμα της Ελλάδας (αναθεώρηση 2019)

[55] Χάρτης Θεμελιωδών Δικαιωμάτων της Ευρώπης, Επίσημη Εφημερίδα της Ευρωπαϊκής Ένωσης (EL), 2012/C 326/02

[56] Dr Λίλιαν Μήτρου (2022), «Πνευματική Ιδιοκτησία», Σημειώσεις Μαθήματος Δίκαιο της Πληροφορικής, Μεταπτυχιακό στην Ασφάλεια Πληροφοριακών και Επικοινωνιακών Συστημάτων, Πανεπιστήμιο Αιγαίου

[57] Dr Ακριβή Βλάχου (2022), «Υδατογράφημα στις Βάσεις Δεδομένων», Σημειώσεις Μαθήματος Ασφάλεια Συστημάτων Βάσεων Δεδομένων, Μεταπτυχιακό στην Ασφάλεια Πληροφοριακών και Επικοινωνιακών Συστημάτων, Πανεπιστήμιο Αιγαίου

[58] Y. Li, H. Guo, and S. Jajodia (2004), "Tamper detection and localization for categorical data using fragile watermarks". In Proceedings of the Digital Rights Management Workshop, pp.73-82.

[60] Πηγή με αριθμητικά σύνολα δεδομένων : https://www.kaggle.com/datasets [πρόσβαση 1/11/2022]

[61] Huiping Guoa YingjiuLib AnyiLiua SushilJajodiaa (2005), "A fragile watermarking scheme for detecting malicious modifications of database relations"

[62] Aihab Khan and Syed Afaq Husain (2013), "A Fragile Zero Watermarking Scheme to Detect and Characterize Malicious Modifications in Database Relations", The Scientific World Journal Volume 2013, Article ID 796726

[63] Jie Guo, "Fragile Watermarking Scheme for Tamper Detectionof Relational Database"

[64] Saeed Arif Shah, Imran Ali Khan2, Syed Zaki Hassan Kazmi, Fariza Hanum Binti Md Nasaruddin, "SEMI-FRAGILE WATERMARKING SCHEME FOR RELATIONAL DATABASE TAMPER DETECTION"

[65] Lancine Camara, Junyi Li, Renfa Li and Wenyong Xie1, "Distortion-Free Watermarking Approach for Relational Database Integrity Checking", Mathematical Problems in Engineering, Volume 2014, Article ID 697165

[66] Yingjiu LI, Vipin Swarup, Sushil Jajodia (2005), «Fingerprinting relational databases: Schemes and specialities», Singapore Management University

[67] Tianxi Ji, Emre Yilmaz, Erman Ayday, Pan Li "The Curse of Correlations for Robust Fingerprinting of Relational Databases"

[68] Tianxi Ji, Emre Yilmaz, Erman Ayday, Pan Li (2022), «Privacy-Preserving Database Fingerprinting»

[69] Tianxi Ji, Erman Ayday, Emre Yilmaz, and Pan Li, "Towards Robust Fingerprinting of Relational Databases by Mitigating Correlation Attacks", IEEE Transactions on Dependable and Secure Computing

[70] Ersin Uzun, Bryan Stephenson (2008), "Security of Relational Databases in Business Outsourcing" HP Laboratories

[71] Ming Zhou, Jianmin Wang, Chaokun Wang and Deyi Li (2007), "A Novel Fingerprinting Architecture for Relational Data", 2007 Inaugural IEEE International Conference on Digital Ecosystems and Technologies

[72] Arti Mohanpurkar, Madhuri Joshi (2015), "Fingerprinting Numeric Databases with Information Preservation and Collusion Avoidance", International Journal of Computer Applications Volume 130 – No.5, November 2015

[73] CUI Xinchun1, QIN Xiaolin2, SHENG Gang3 "A Weighted Algorithm for Watermarking Relational Databases"

[74] Βασιλείου, Δημητρίου "Πειραματική εφαρμογή client-server αναγνώρισης δακτυλικών αποτυπωμάτων", Π.Μ.Σ. στην Επιστήμη και Τεχνολογία Υπολογιστών Κατεύθυνση: Πληροφορική και Εφαρμογές

[75] E.Solami, M.Kamran , M.Alkatheiri, F.Rafiq and Ahmed.Alghamdi (2020), "Fingerprinting of Relational Databases for Stopping the Data Theft", MDPI Electronics

[76] Ming-Ru Xie, Chia-Chun Wu, Jau-Ji Shen, Min-Shiang Hwang (2016), "A Survey of Data Distortion Watermarking Relational Databases", International Journal of Network Security, Vol.18, No.6, PP.1022-1033, Nov. 2016

[77] Alapan Arnab and Andrew Hutchison, "REQUIREMENT ANALYSIS OF DRM SYSTEMS"

[78] H.L. Jonker1, S. Mauw1, J.H.S. Verschuren, A.T.S.C. Schoonen, "SECURITY ASPECTS OF DRM SYSTEMS"

[79] Digital Rights Management (DRM) https://www.fortinet.com/resources/cyberglossary/digital-rights-management-drm [πρόσβαση 8/10/2022]

[80] "What is Digital Rights Management (DRM) and How Does it Work For You" https://www.vitrium.com/what-is-digital-rights-management-drm-and-how-does-it-work-for-you [πρόσβαση 8/10/2022]

[81] "Digital Rights Management (DRM) Architectures" https://www.dlib.org/dlib/june01/iannella/06iannella.html [πρόσβαση 1/10/2022]

[82] "A Survey of Digital Rights Management Technologies" https://www.cse.wustl.edu/~jain/cse571-11/ftp/drm/ [πρόσβαση 1/10/2022]

[83] el.wikipedia.org/wiki/Κρυπτογραφία [πρόσβαση 23/10/2022]

[84] el.wikipedia.org/wiki/Στεγανογραφία [πρόσβαση 23/10/2022]

[85] el.wikipedia.org/wiki/Υδατογράφημα [πρόσβαση 23/10/2022]

[86] Saman Iftikhar, M Kamran, Zahid Anwar, "A survey on reversible watermarking techniques for relational databases», January 2015, Wiley Online Library, Security Communications and Networks 2015

[87] Manoj Kumar, Om Prakash Verma, "A Robust Database Watermarking Algorithm using Bezier Cubic Curves", International Journal of Pure and Applied Mathematics, Volume 117 No. 20 2017, 881-886

[88] Manoj Kumar, O. P. Verma, Archana Saxena, "Elliptic Curve Cryptography (ECC) based Relational Database Watermarking", International Journal of Computer Applications (0975 – 8887), Volume 154 – No.6, November 2016

[89] Raju Halder, Agostino Cortesi, "Persistent Watermarking of Relational Databases", 2010 International Conference on Advances in Communication, Network, and Computing

[90] Yingjiu LI, Robert H. DENG, "Publicly Verififiable Ownership Protection for Relational Databases", ASIACCS '06: Proceedings of the ACM Symposium on Information, Computer and Communications Security: Taipei, Taiwan, 21-24 March. 78-89.

[91] Ms. Vrushali B. Patil1, Prof. P. M. Yawalkar, "A Review on Multi-attribute Watermarking Technique for Relational Data", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 4, Issue 12, December 2015

[92] Mahmoud E. Farfoura, Shi-Jinn Horng, Jui-Lin Lai, Ray-Shine Run, Rong-Jian Chen, Muhammad Khurram Khan, "A blind reversible method for watermarking relational databases based on a time-stamping protocol", Expert Systems with Applications 39 (2012) 3185–3196

[93] Javier Franco Contreras and Gouenou Coatrieux, "Protection of Relational Databases by Means of Watermarking: Recent Advances and Challenges", Advances in Security in Computing and Communications, July 19th, 2017

Παράρτημα Ι : Τα αναλυτικά σενάρια και τα αποτελέσματα των πειραμάτων

Όλα τα σενάρια με παραλλαγές στις παραμέτρους τους, δημιουργήθηκαν με αντίστοιχα scripts που καλούν το βασικό παραμετρικό πρόγραμμα, συλλέγουν τα αποτελέσματα ανίχνευσης και παράγουν τα γραφήματα.

7.5 Τα Σενάρια

Σε αυτό το παράρτημα βρίσκονται όλα τα scripts, τα αποτελέσματα και τα διαγράμματα των πειραμάτων που έγιναν. Τα παρακάτω σενάρια εκτελέστηκαν σε 3 διαφορετικά datasets. Οι τιμές των ανεξάρτητων μεταβλητών με κόκκινο δοκιμάστηκαν σε όλους τους συνδυασμούς.

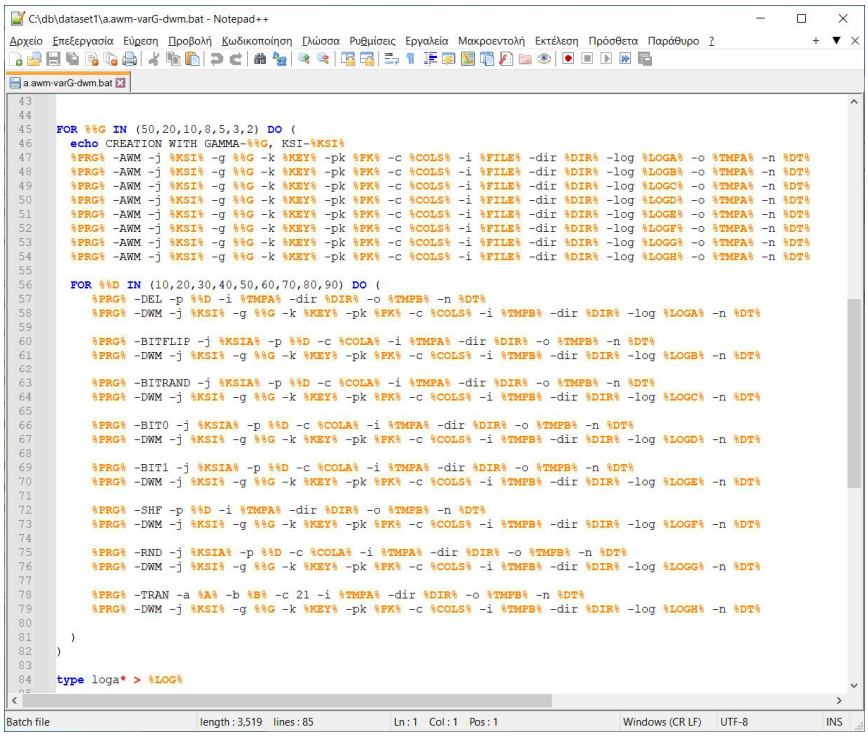
Ομάδα	ΣΕΝΑΡΙΑ (Ακολουθίες ενεργειών)	Συνδυασμοί Μεταβλητών
A	1.Add WM -> Subset Deletion Attack -> Detect WM 2.Add WM -> Bit Flip Attack -> Detect WM 3.Add WM -> Bit Rnd Attack -> Detect WM 4.Add WM -> Shuffle Attack -> Detect WM	**Σταθερές : ξ=5, ν=4** **Συνδυασμοί Μεταβλητών** **γ=50,20,10,8,5,4,2** **%επίθεσης=10,20, … , 90**
B	1. Add WM -> Subset Deletion Attack -> Detect WM 2.Add WM -> Bit Flip Attack -> Detect WM 3.Add WM -> Bit Rnd Attack -> Detect WM 4.Add WM -> Shuffle Attack -> Detect WM	**Σταθερές : %επίθεσης=20%, ξ2=6** **Συνδυασμοί Μεταβλητών** **γ=50,20,10,8,5,3** **ξ=4,5,6,7,8,9**
C	1. Add WM -> Subset Deletion Attack -> Detect WM 2.Add WM -> Bit Flip Attack -> Detect WM 3.Add WM -> Bit Rnd Attack -> Detect WM 4.Add WM -> Shuffle Attack -> Detect WM	**Σταθερές : γ=5, ν=4, ξ2=8** **Συνδυασμοί Μεταβλητών** **ξ=3,4,5,6,7,8** **%επίθεσης=10,20, … , 90**
D	Add WM -> CFO Attack -> Detect WM	**Σταθερές : ξ=6, ν=4** **γ=50,20,10,8,5,4,2**

Ομάδα	ΣΕΝΑΡΙΑ (Ακολουθίες ενεργειών)	Συνδυασμοί Μεταβλητών
		ξ2=2,3,4,5 **γ2=5,10,15,20,25,30,35,40,45**
E	Add WM -> Subset Deletion Attack -> Detect WM Add WM -> Bit 0 Attack -> Detect WM Add WM -> Bit 1 Attack -> Detect WM Add WM -> Bit Flip Attack -> Detect WM Add WM -> Bit Rand Attack -> Detect WM Add WM -> Rounding -> Detect WM Add WM -> Transformation -> Detect WM Add WM -> Shuffle Attack -> Detect WM Add WM -> CFO Attack -> Detect WM (ξ2=4)	**Σταθερές : ξ=5** **Συνδυασμοί Μεταβλητών** **γ=20,10,5** **%επίθεσης=10,20, … , 90**
F	Add WM -> Deletion -> Bit Flip (ξ2=7) -> Shuffle -> Detect WM (Συνδυασμός πολλαπλών επιθέσεων)	**Σταθερές : ξ=6, ν=4** **Συνδυασμοί Μεταβλητών** **γ=50,20,10,8,5,4,2** **%επίθεσης=10,20, … , 90**
G	Add WM -> CFO Attack -> Detect WM	**Σταθερές : γ=20, ν=4** **Συνδυασμοί Μεταβλητών** **ξ=3,4,5,6,7,8,9,10** **γ2=8,16,24,32** **ξ2=2,3,4,5,6,7,8,9,10**
H	Add WM **weight/attribute** -> BitFlipAttack -> Detect WM	**Σταθερές : ξ=8, ν=4** **Βαρύτητες πιθανότητας επιλογής χαρακτηριστικών [10,6,3,1]** **Συνδυασμοί Μεταβλητών** **γ=50,20,10,8,5,4,2** **%επίθεσης=10,20, … , 90**
I	Add WM **ξ/attribute** -> BitFlipAttack -> Detect WM	**Σταθερές : ν=4,** **ξ/χαρακτηριστικό [8,4,2,1]** **Συνδυασμοί Μεταβλητών** **γ=50,20,10,8,5,4,2** **%επίθεσης=10,20, … , 90**
J	Brute Force key guessing of watermark	**Σταθερά: creation key=gHUI!216** **Δοκιμές με** **detection key= gHUI!21{N}** **{N}=1,2 … 20**

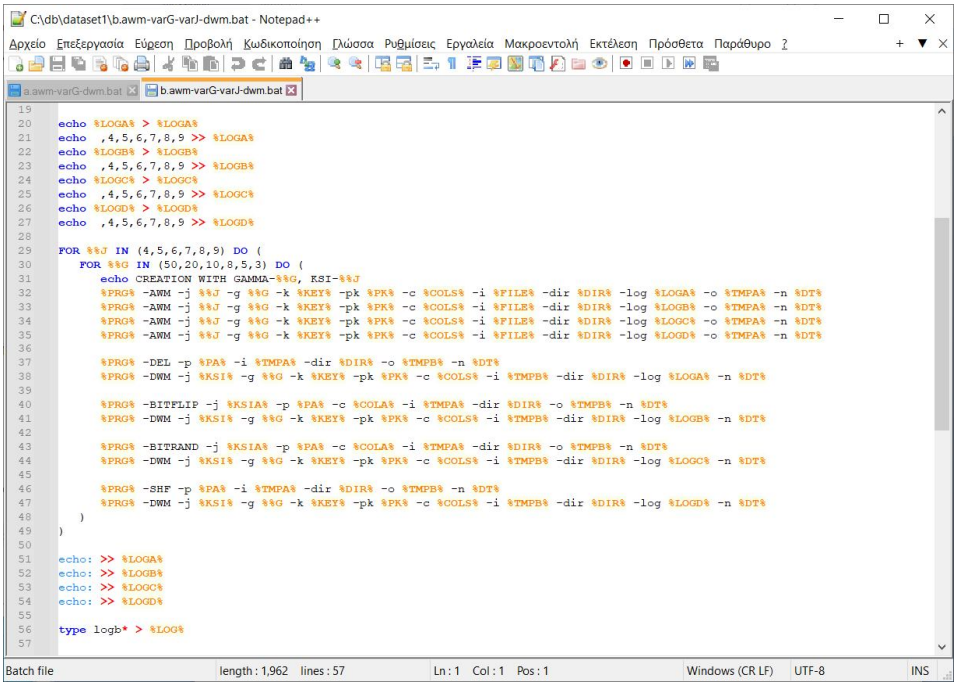
7.6 Οι δέσμες ενεργειών που υλοποιούν τα σενάρια

Οι δέσμες ενεργειών των 10 σεναρίων που χρησιμοποιήθηκαν στα 3 dataset αποτελούν αναπόσπαστο κομμάτι της εργασίας. Τα scripts εκτελέστηκαν στα 3 datasets με διαφοροποίηση των παραμέτρων που αφορούν το βασικό κλειδί, τα υποψήφια χαρακτηριστικά για μαρκάρισμα και τις τιμές ξ. Όλα τα διαγράμματα των πειραμάτων επίσης αποτελούν μέρος της εργασίας.

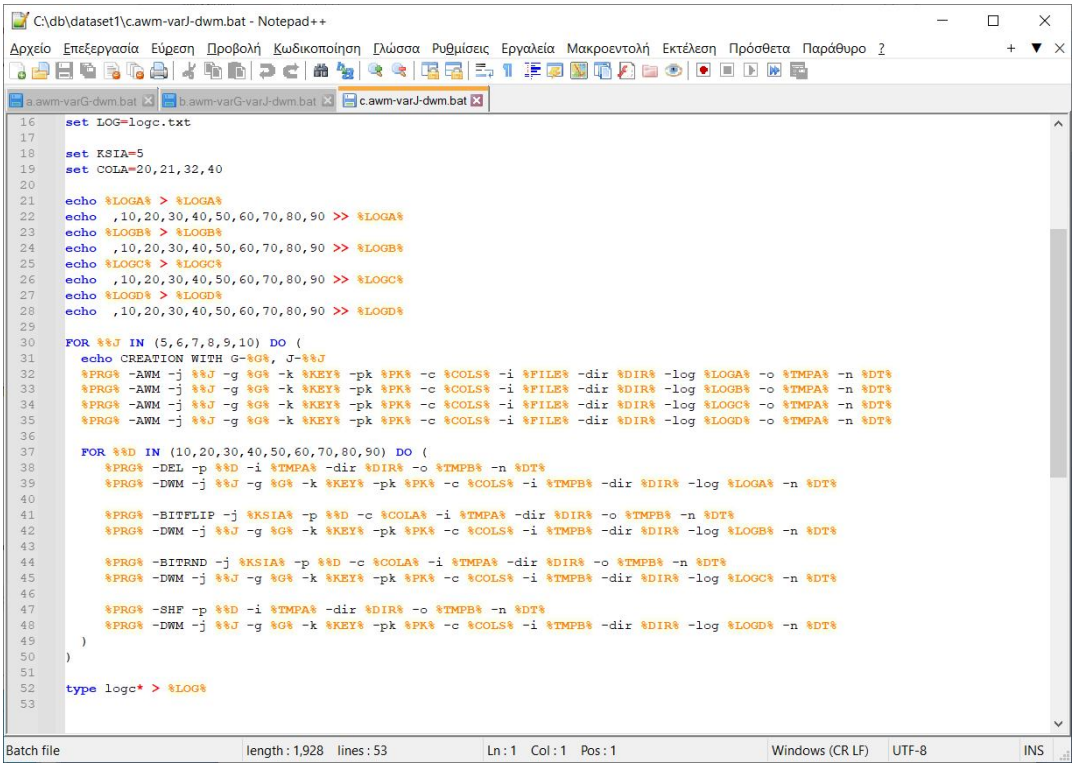
7.6.1 Σενάρια ομάδας Α, μεταβαλλόμενου γ και ποσοστού επίθεσης (με σταθερό ξ)



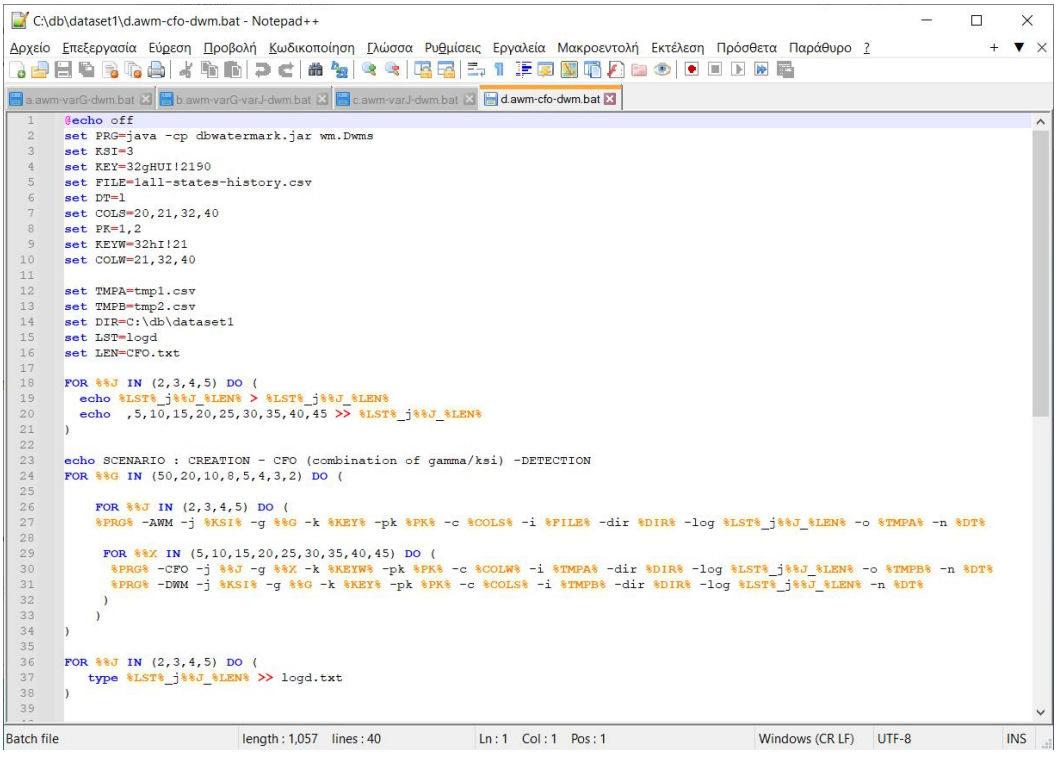
7.6.2 Σενάρια ομάδας Β, μεταβαλλόμενου γ και μεταβαλλόμενου ξ (με σταθερό %επίθεσης)



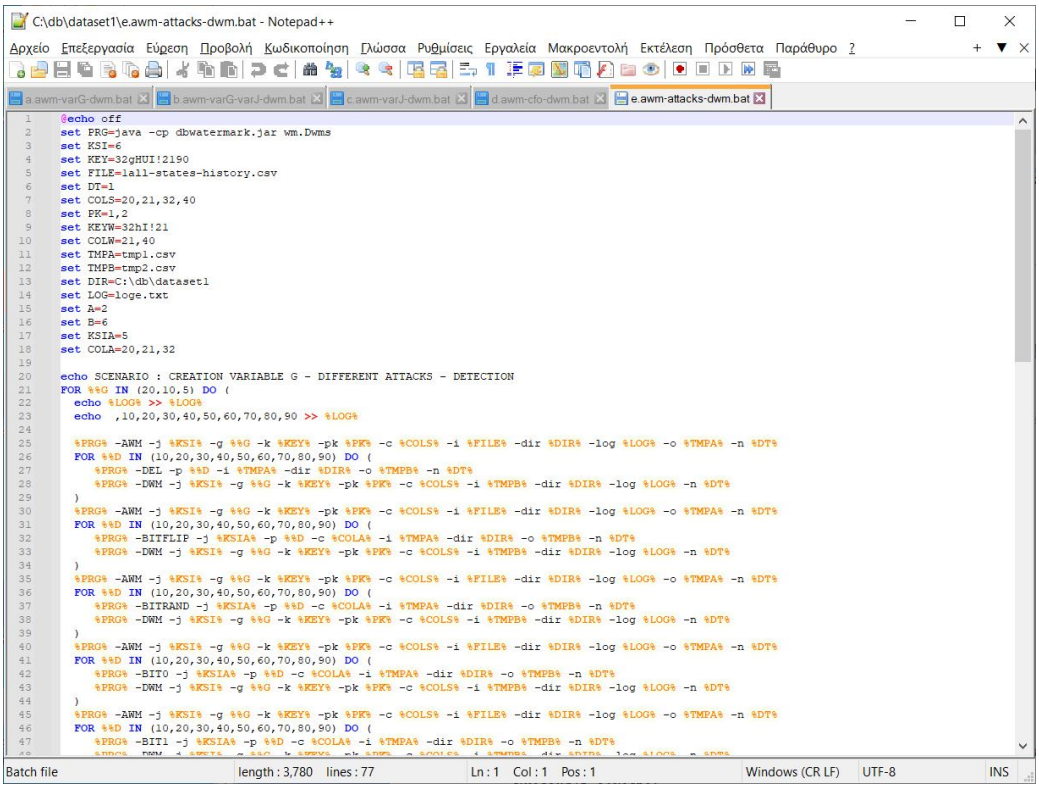
7.6.3 Σενάρια ομάδας C, μεταβαλλόμενου ξ και ποσοστού επίθεσης (με σταθερό γ)



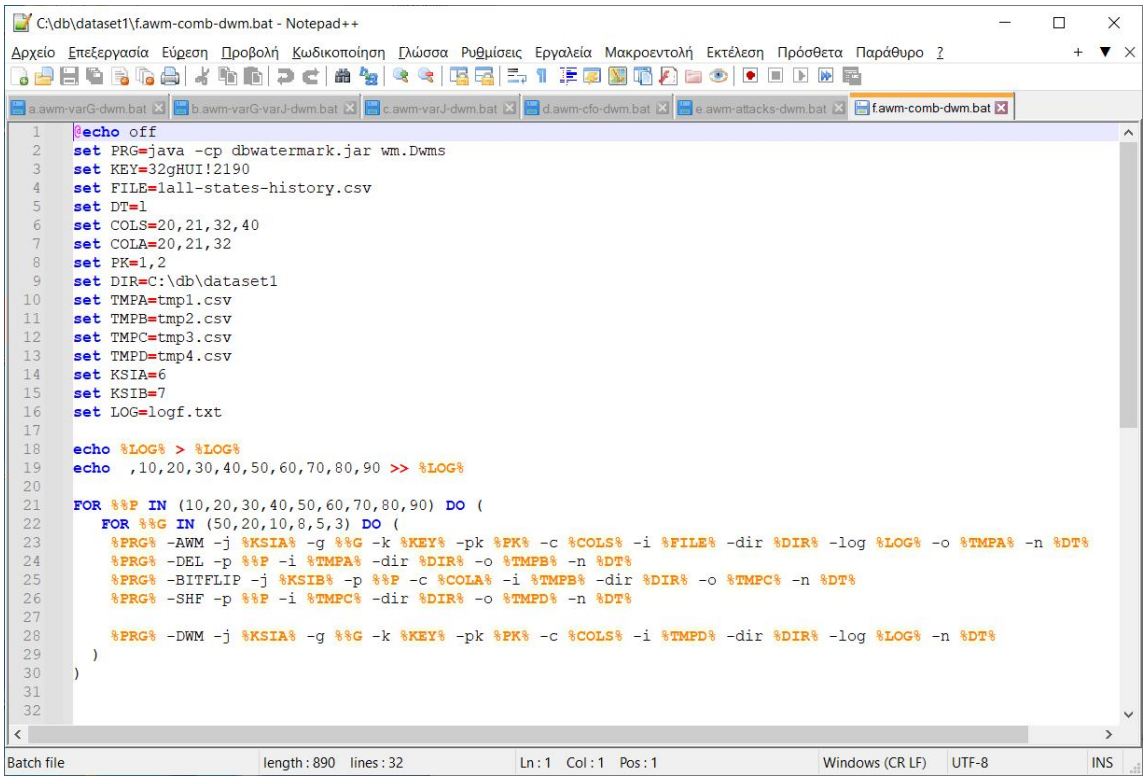
7.6.4 Σενάρια ομάδας D (CFO), μεταβαλλόμενων γ, ξ2, γ2 (με σταθερό ξ)



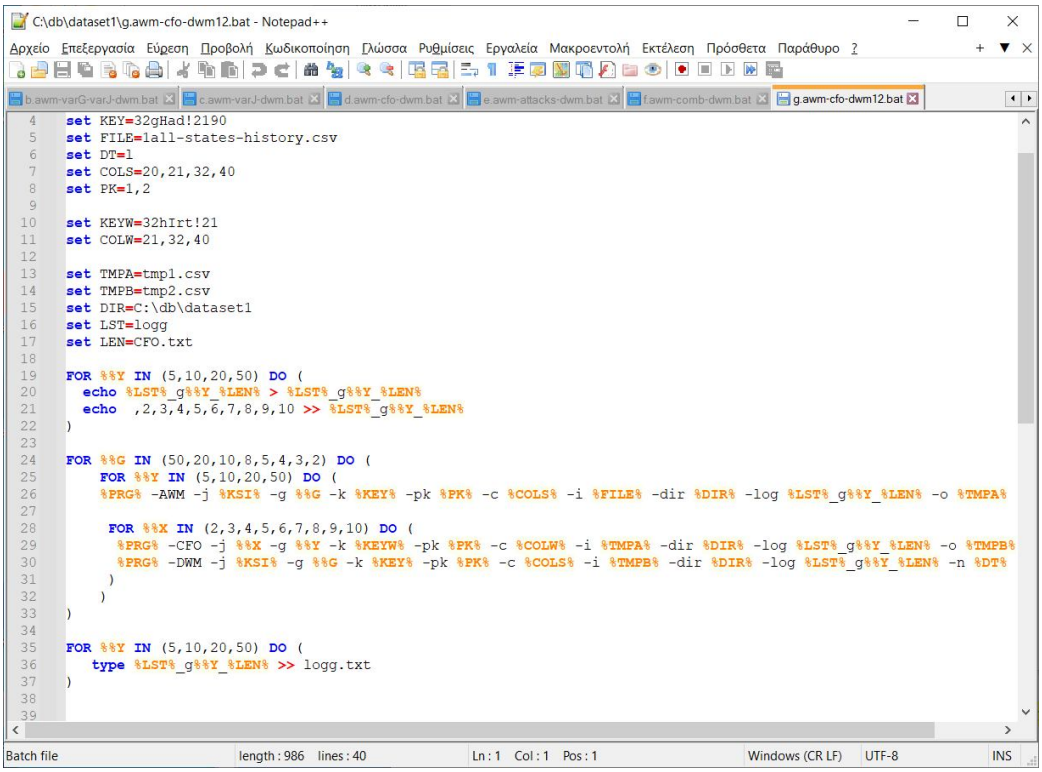
7.6.5 Σενάρια ομάδας E, μεταβαλλόμενου τύπου επιθέσεων γ και ποσοστού επίθεσης (με σταθερό ξ)



7.6.6 Σενάρια ομάδας F, πολλαπλών επιθέσεων, μεταβαλλόμενου γ και ποσοστού επίθεσης (με σταθερό ξ)

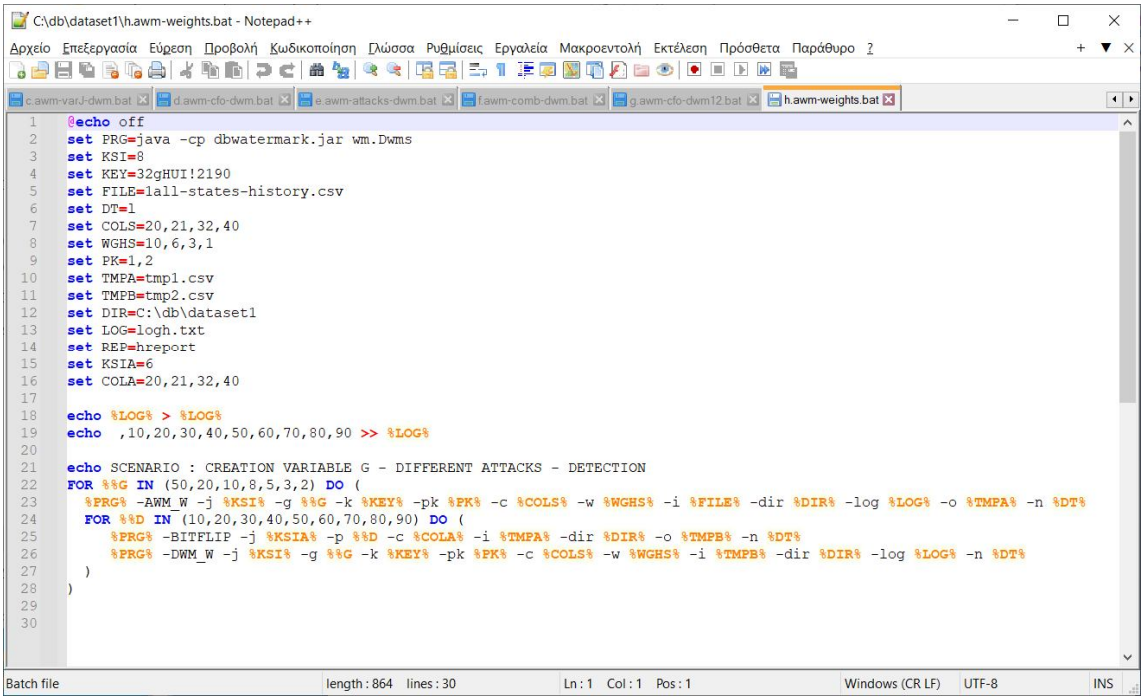


7.6.7 Σενάρια ομάδας G (CFO), μεταβαλλόμενων ξ, ξ2, γ2 (με σταθερό γ)



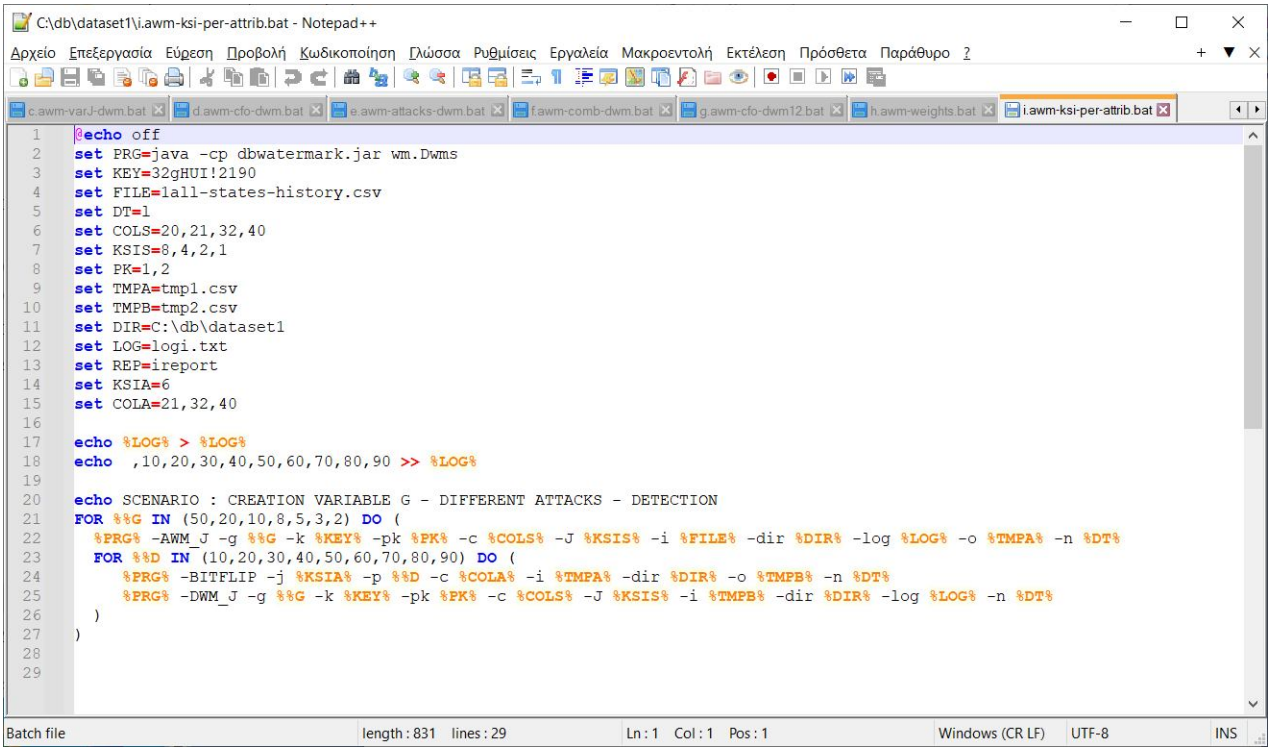
```
C:\db\dataset1\g.awm-cfo-dwm12.bat - Notepad++
Αρχείο Επεξεργασία Εύρεση Προβολή Κωδικοποίηση Γλώσσα Ρυθμίσεις Εργαλεία Μακροεντολή Εκτέλεση Πρόσθετα Παράθυρο ?
b.awm-varG-varJ-dwm.bat | c.awm-varJ-dwm.bat | d.awm-cfo-dwm.bat | e.awm-attacks-dwm.bat | f.awm-comb-dwm.bat | g.awm-cfo-dwm12.bat
 4  set KEY=32gHad!2190
 5  set FILE=1all-states-history.csv
 6  set DT=1
 7  set COLS=20,21,32,40
 8  set PK=1,2
 9
10  set KEYW=32hIrt!21
11  set COLW=21,32,40
12
13  set TMPA=tmp1.csv
14  set TMPB=tmp2.csv
15  set DIR=C:\db\dataset1
16  set LST=logg
17  set LEN=CFO.txt
18
19  FOR %%Y IN (5,10,20,50) DO (
20    echo %LST%_g%%Y_%LEN% > %LST%_g%%Y_%LEN%
21    echo  ,2,3,4,5,6,7,8,9,10 >> %LST%_g%%Y_%LEN%
22  )
23
24  FOR %%G IN (50,20,10,8,5,4,3,2) DO (
25      FOR %%Y IN (5,10,20,50) DO (
26      %PRG% -AWM -j %KSI% -g %%G -k %KEY% -pk %PK% -c %COLS% -i %FILE% -dir %DIR% -log %LST%_g%%Y_%LEN% -o %TMPA%
27
28       FOR %%X IN (2,3,4,5,6,7,8,9,10) DO (
29        %PRG% -CFO -j %%X -g %%Y -k %KEYW% -pk %PK% -c %COLW% -i %TMPA% -dir %DIR% -log %LST%_g%%Y_%LEN% -o %TMPB%
30        %PRG% -DWM -j %KSI% -g %%G -k %KEY% -pk %PK% -c %COLS% -i %TMPB% -dir %DIR% -log %LST%_g%%Y_%LEN% -n %DT%
31       )
32      )
33  )
34
35  FOR %%Y IN (5,10,20,50) DO (
36     type %LST%_g%%Y_%LEN% >> logg.txt
37  )
38
39
Batch file | length : 986 lines : 40 | Ln : 1 Col : 1 Pos : 1 | Windows (CR LF) | UTF-8 | INS
```

7.6.8 Σενάρια ομάδας H, με διαφορετικές βαρύτητες στα χαρακτηριστικά, μεταβαλλόμενου γ και ποσοστού επίθεσης (με σταθερά ξ και ξ2)

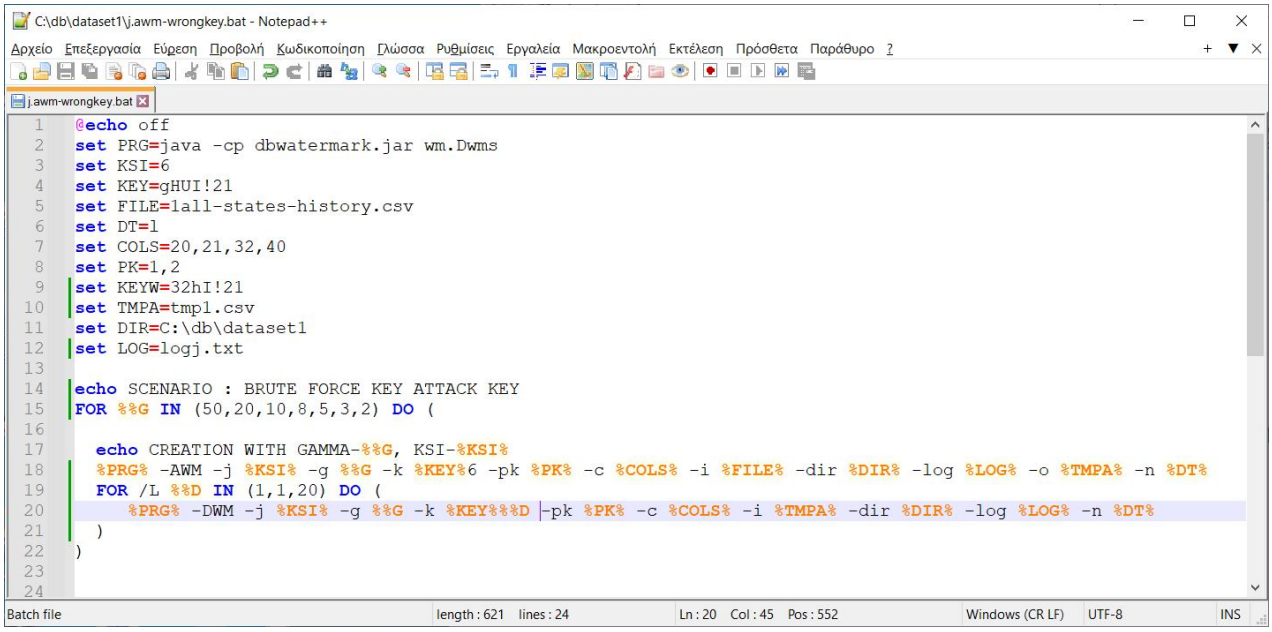


```
C:\db\dataset1\h.awm-weights.bat - Notepad++
Αρχείο Επεξεργασία Εύρεση Προβολή Κωδικοποίηση Γλώσσα Ρυθμίσεις Εργαλεία Μακροεντολή Εκτέλεση Πρόσθετα Παράθυρο ?
c.awm-varJ-dwm.bat | d.awm-cfo-dwm.bat | e.awm-attacks-dwm.bat | f.awm-comb-dwm.bat | g.awm-cfo-dwm12.bat | h.awm-weights.bat
 1  @echo off
 2  set PRG=java -cp dbwatermark.jar wm.Dwms
 3  set KSI=8
 4  set KEY=32gHUI!2190
 5  set FILE=1all-states-history.csv
 6  set DT=1
 7  set COLS=20,21,32,40
 8  set WGHS=10,6,3,1
 9  set PK=1,2
10  set TMPA=tmp1.csv
11  set TMPB=tmp2.csv
12  set DIR=C:\db\dataset1
13  set LOG=logh.txt
14  set REP=hreport
15  set KSIA=6
16  set COLA=20,21,32,40
17
18  echo %LOG% > %LOG%
19  echo  ,10,20,30,40,50,60,70,80,90 >> %LOG%
20
21  echo SCENARIO : CREATION VARIABLE G - DIFFERENT ATTACKS - DETECTION
22  FOR %%G IN (50,20,10,8,5,3,2) DO (
23    %PRG% -AWM_W -j %KSI% -g %%G -k %KEY% -pk %PK% -c %COLS% -w %WGHS% -i %FILE% -dir %DIR% -log %LOG% -o %TMPA% -n %DT%
24    FOR %%D IN (10,20,30,40,50,60,70,80,90) DO (
25       %PRG% -BITFLIP -j %KSIA% -p %%D -c %COLA% -i %TMPA% -dir %DIR% -o %TMPB% -n %DT%
26       %PRG% -DWM_W -j %KSI% -g %%G -k %KEY% -pk %PK% -c %COLS% -w %WGHS% -i %TMPB% -dir %DIR% -log %LOG% -n %DT%
27    )
28  )
29
30
Batch file | length : 864 lines : 30 | Ln : 1 Col : 1 Pos : 1 | Windows (CR LF) | UTF-8 | INS
```

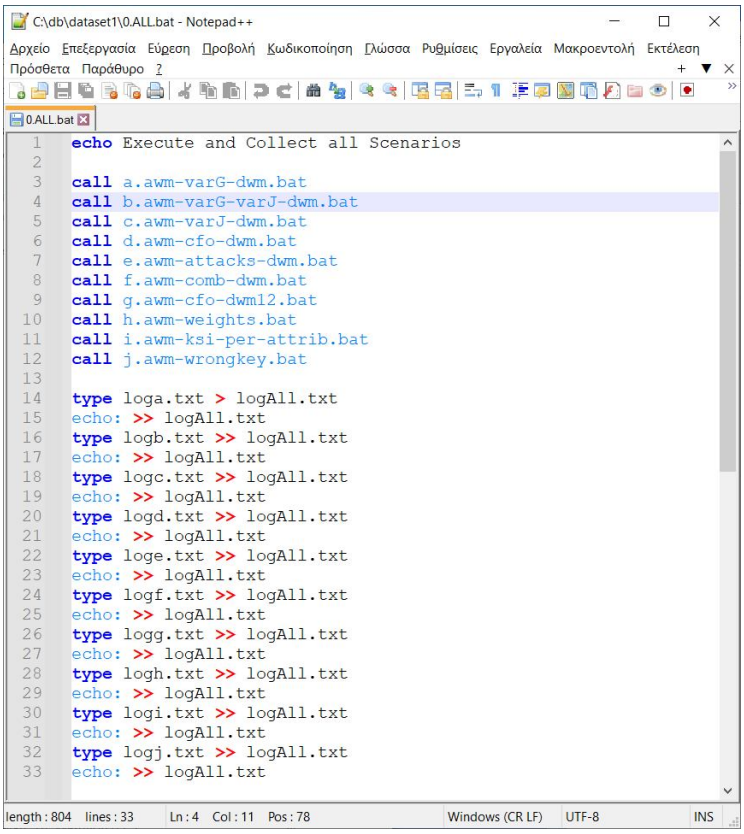
7.6.9 Σενάρια ομάδας I με διαφορετικά ξ ανά ιδιότητα, μεταβαλλόμενου γ και ποσοστού επίθεσης (με σταθερό ξ2)



7.6.10 Σενάρια ομάδας J με επιθέσεις τύπου [brute force] μέσα από μικρές διαφοροποιήσεις του κλειδιού ανίχνευσης (με σταθερό ξ)

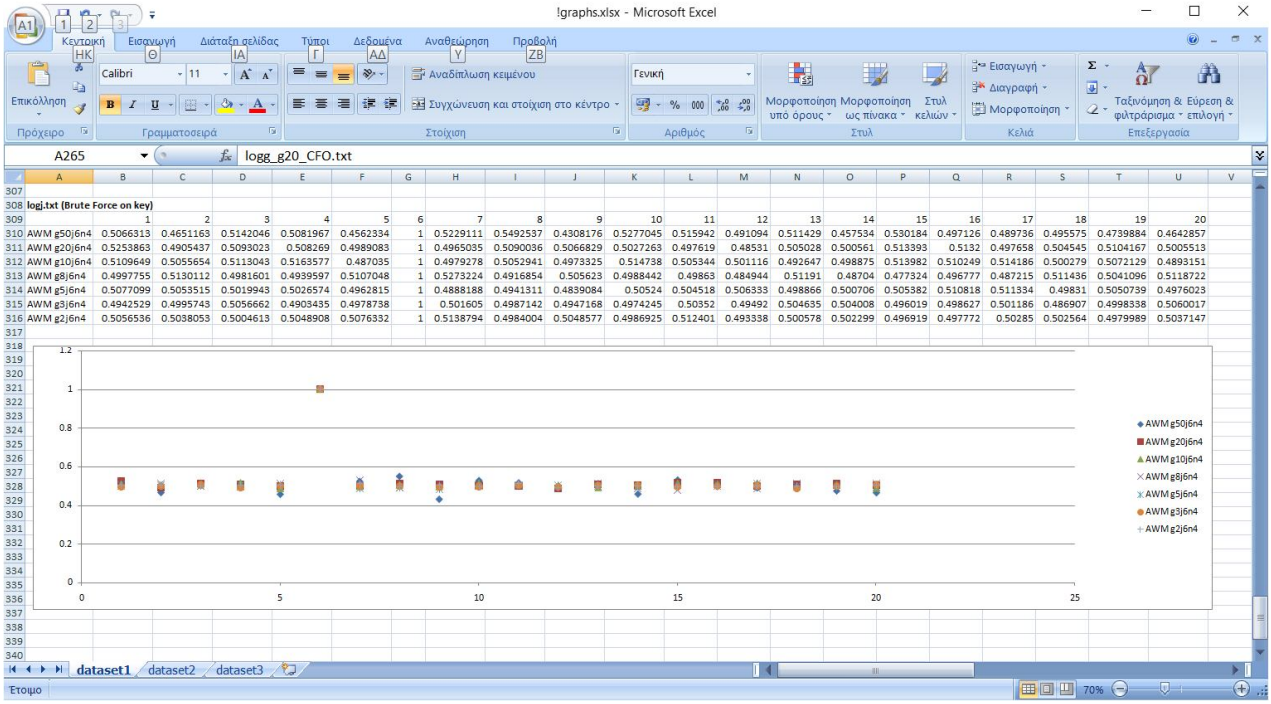


7.6.11 Αυτοματοποιημένη εκτέλεση όλων των σεναρίων των δέκα ομάδων A-J

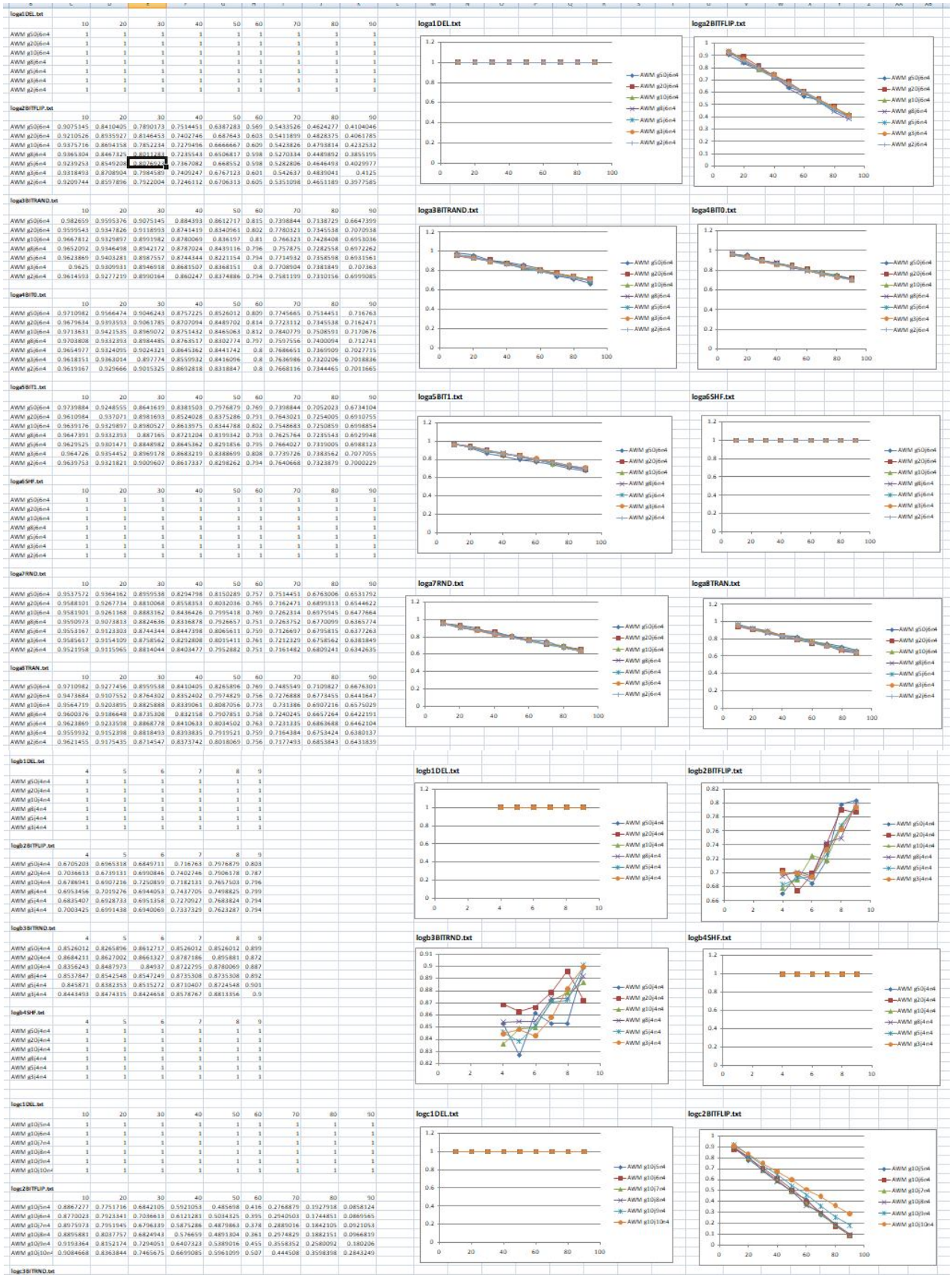


7.6.12 Σενάριο Brute Force Attack με δοκιμές στον κωδικό ανίχνευσης

Το σωστό κλειδί ανιχνεύει μια υδατογράφηση (χωρίς επιθέσεις) σε βαθμό 100%. Τα λάθος κλειδιά οδηγούν σε ποσοστά επιτυχίας που κυμαίνονται κοντά στο 50%.



7.7 Εφαρμογή σεναρίων στο dataset 1: «Περιστατικά covid USA»



logc3BITRND.txt

	10	20	30	40	50	60	70	80	90
AWM g10j5n4	0.8867277	0.7751716	0.6842105	0.5921053	0.485698	0.416	0.2768879	0.1927918	0.0858124
AWM g10j6n4	0.8770023	0.7923341	0.7036613	0.6121281	0.5034325	0.395	0.2940503	0.1744851	0.0869565
AWM g10j7n4	0.8975973	0.7951945	0.6796339	0.5875286	0.4879863	0.378	0.2689016	0.1842105	0.0921053
AWM g10j8n4	0.8895881	0.8057757	0.6824943	0.576659	0.4891304	0.361	0.2974829	0.1882151	0.0966819
AWM g10j9n4	0.9193364	0.8152174	0.7294053	0.6407323	0.5389016	0.455	0.3558352	0.2580092	0.180206
AWM g10j10n4	0.9084668	0.8363844	0.7465675	0.6899085	0.5961099	0.507	0.444508	0.3598398	0.2843249

logc4SHF.txt

	10	20	30	40	50	60	70	80	90
AWM g10j5n4	1	1	1	1	1	1	1	1	1
AWM g10j6n4	1	1	1	1	1	1	1	1	1
AWM g10j7n4	1	1	1	1	1	1	1	1	1
AWM g10j8n4	1	1	1	1	1	1	1	1	1
AWM g10j9n4	1	1	1	1	1	1	1	1	1
AWM g10j10n4	1	1	1	1	1	1	1	1	1

logd_j2_CFO.txt

	5	10	15	20	25	30	35	40	45
AWM g50j3n4	0.9942197	0.9971098	0.9971098	1	0.9942197	1	1	1	1
AWM g20j3n4	0.9908467	0.993135	0.9965675	0.9954233	0.9954233	0.998	1	0.9988558	0.9988558
AWM g10j3n4	0.9919817	0.9959908	0.9965636	0.9971363	0.997709	0.998	0.9988545	0.9988545	0.9982818
AWM g8j3n4	0.9929478	0.996709	0.9976493	0.9981194	0.9995298	0.999	1	0.9990597	0.9990597
AWM g5j3n4	0.9929299	0.9968891	0.9971739	0.9974548	0.9988688	0.999	0.9994344	0.9988688	0.9988688
AWM g4j3n4	0.9918529	0.9946462	0.9969739	0.9969739	0.9983706	0.998	0.9995344	0.9983706	0.9983378
AWM g3j3n4	0.9909247	0.9953767	0.9974315	0.997774	0.997774	0.999	0.9986302	0.9988014	0.9994863
AWM g2j3n4	0.9921085	0.9956542	0.9970265	0.9975984	0.9985133	0.998	0.9990851	0.9988564	0.9985133

logd_j3_CFO.txt

	5	10	15	20	25	30	35	40	45
AWM g50j3n4	0.9942197	1	0.9942197	1	0.9971098	1	1	1	1
AWM g20j3n4	0.993135	0.9954233	0.9965675	0.9977117	0.9988558	0.999	1	1	1
AWM g10j3n4	0.9954181	0.9965636	0.9982818	0.9988545	0.9994273	0.999	0.9994273	1	1
AWM g8j3n4	0.990127	0.9948284	0.9962388	0.996709	0.9981194	0.997	0.9985896	0.9985896	0.9990597
AWM g5j3n4	0.9949095	0.9974548	0.9988688	0.9991516	0.9997172	1	0.9994344	1	0.9997172
AWM g4j3n4	0.9909218	0.9955773	0.9965084	0.9976723	0.9983378	0.998	0.9986033	0.9990689	0.9986033
AWM g3j3n4	0.989726	0.994863	0.9969178	0.9974315	0.9981164	0.998	0.9981164	0.9991438	0.9984589
AWM g2j3n4	0.9918801	0.9956542	0.9967978	0.9979414	0.9985133	0.998	0.9988564	0.9991995	0.9985133

logd_j4_CFO.txt

	5	10	15	20	25	30	35	40	45
AWM g50j3n4	1	1	1	1	1	1	1	1	1
AWM g20j3n4	0.9919909	0.9942792	0.9977117	0.9954233	0.9988558	0.998	1	0.9977117	0.9988558
AWM g10j3n4	0.9919817	0.9959908	0.9971363	0.9971363	0.9994273	0.998	0.9988545	0.9982818	0.9982818
AWM g8j3n4	0.996709	0.9985896	0.9985896	0.9990597	1	0.999	0.9990597	0.9995298	0.9995298
AWM g5j3n4	0.9943439	0.9974548	0.9983032	0.9980204	0.9997172	0.999	0.9994344	0.9988688	0.9988688
AWM g4j3n4	0.9944134	0.9969739	0.9974395	0.9983706	0.9990689	0.998	0.9995344	0.9990689	0.9986033
AWM g3j3n4	0.9936644	0.997774	0.9982877	0.9989726	0.9986302	0.999	0.9993151	0.9993151	0.9994863
AWM g2j3n4	0.9941675	0.9967978	0.9977127	0.9983989	0.9989708	0.999	0.9993138	0.9990851	0.9989708

logd_j5_CFO.txt

	5	10	15	20	25	30	35	40	45
AWM g50j3n4	0.9942197	0.9942197	1	0.9971098	0.9971098	1	1	1	1
AWM g20j3n4	0.9942792	0.9954233	1	0.9965675	0.9977117	1	1	1	1
AWM g10j3n4	0.9954181	0.9965636	1	0.9982818	0.9988545	1	1	1	1
AWM g8j3n4	0.9943582	0.9976493	0.9985896	0.9985896	0.9995298	0.999	1	1	0.9990597
AWM g5j3n4	0.9954753	0.9971719	0.9994344	0.998586	0.9991516	1	0.9997172	0.9997172	0.9997172
AWM g4j3n4	0.9934823	0.9962756	0.997905	0.9983706	0.9988361	0.999	0.9997672	0.9993017	0.9986033
AWM g3j3n4	0.9936644	0.9967466	0.9993151	0.9982877	0.9984589	0.999	0.9989726	0.9991438	0.9996575
AWM g2j3n4	0.9956542	0.9977127	0.9989708	0.9989708	0.9993138	0.999	0.9997713	0.9994282	0.9993138

loge.txt

	10	20	30	40	50	60	70	80	90
DEL g20j6n4	1	1	1	1	1	1	1	1	1
BITFLIP	0.9382151	0.8707094	0.8169336	0.7322655	0.6624714	0.6	0.5320366	0.4702517	0.4187643
BITRAND	0.9553776	0.9347826	0.8935927	0.8695652	0.832952	0.805	0.7643023	0.7025172	0.6922197
BIT0	0.9622425	0.9290618	0.8924485	0.8707094	0.8443936	0.802	0.7826087	0.7368421	0.7139588
BIT1	0.9530892	0.9153318	0.8947368	0.8501144	0.8020595	0.797	0.7459954	0.7299771	0.6899313
SHF	1	1	1	1	1	1	1	1	1
RND	0.9530892	0.9130435	0.8764302	0.8363844	0.812357	0.769	0.7185355	0.694508	0.6544622
TRAN	0.9599543	0.9336385	0.8798627	0.8455378	0.8135011	0.763	0.7219679	0.6819222	0.6407323
CFO	0.9977117	0.9977117	0.9988558	0.9988558	1	0.999	1	0.9988558	1

loge.txt

	10	20	30	40	50	60	70	80	90
DEL g10j6n4	1	1	1	1	1	1	1	1	1
BITFLIP	0.9272628	0.851661	0.7852234	0.7416953	0.6758305	0.613	0.5395189	0.489118	0.4209622
BITRAND	0.9633448	0.9284078	0.8946163	0.8636884	0.8367698	0.787	0.7726231	0.7365407	0.7107675
BIT0	0.9656357	0.932437	0.906071	0.8631157	0.847079	0.804	0.7743413	0.7462772	0.7147766
BIT1	0.9633448	0.9306988	0.8917526	0.8636884	0.8344788	0.801	0.7646048	0.7342497	0.69874
SHF	1	1	1	1	1	1	1	1	1

loge.txt

	10	20	30	40	50	60	70	80	90
DEL g5j6n4	1	1	1	1	1	1	1	1	1
BITFLIP	0.9301471	0.8713235	0.8000566	0.7270927	0.6648756	0.598	0.5376131	0.4717194	0.4052602
BITRAND	0.9618213	0.9366516	0.8970588	0.8602941	0.8308824	0.793	0.7794118	0.738405	0.7044683
BIT0	0.969457	0.9324095	0.8857466	0.8614253	0.8365384	0.798	0.7734729	0.7372738	0.7058824
BIT1	0.9618213	0.9290158	0.9018665	0.8628393	0.8297511	0.798	0.7652715	0.7293552	0.696267
SHF	1	1	1	1	1	1	1	1	1
RND	0.9604073	0.9244909	0.8744344	0.8385181	0.8161765	0.769	0.7287896	0.6934389	0.6473416
TRAN	0.9544683	0.9114819	0.8829185	0.8441742	0.8147625	0.768	0.7259616	0.6855204	0.6459276
CFO	0.9971719	0.9983032	0.9991516	0.9988688	0.998586	0.999	1	0.9991516	1

logf.txt

	10	20	30	40	50	60	70	80	90
AWM g50j6n4	0.9032258	0.7931014	0.7317073	0.6451613	0.5865922	0.449	0.3898305	0.3333333	0.3214286
AWM g20j6n4	0.9088542	0.8248588	0.7553366	0.6698293	0.6229117	0.521	0.4514926	0.3368984	0.2985075
AWM g10j6n4	0.9165589	0.8310469	0.7721202	0.6968542	0.6275626	0.526	0.4323233	0.3042169	0.2530121
AWM g8j6n4	0.9051133	0.8178572	0.7638889	0.6556082	0.6083172	0.506	0.44	0.3558897	0.259887
AWM g5j6n4	0.9122639	0.8292857	0.7359194	0.6673068	0.596222	0.509	0.4389535	0.3551829	0.2983051
AWM g3j6n4	0.9166505	0.8339852	0.7514837	0.674626	0.5819925	0.519	0.4636637	0.3655617	0.2969349

logg_g5_CFO.txt

	2	3	4	5	6	7	8	9	10
AWM g50j6n4	0.9971671	1	0.9971671	1	0.9971671	1	0.9971671	1	1
AWM g20j6n4	0.9941928	0.9988385	0.9965157	0.9965157	0.9930314	0.997	0.9976772	0.9965157	0.9965157
AWM g10j6n4	0.9942955	0.9971477	0.9937251	0.9965773	0.9925842	0.998	0.9971477	0.9971477	0.9971477
AWM g8j6n4	0.995616	0.9969312	0.9947392	0.9929856	0.9969312	0.996	0.9960544	0.9995616	0.9960544
AWM g5j6n4	0.9942529	0.9968391	0.9945402	0.9971265	0.9925287	0.997	0.9968391	0.9971265	0.9974138
AWM g4j6n4	0.9960215	0.9968979	0.9960215	0.9951252	0.9962331	0.996	0.997342	0.9977842	0.9968979
AWM g3j6n4	0.996203	0.9972385	0.9968652	0.9955126	0.9958578	0.997	0.9979289	0.9984466	0.9972385
AWM g2j6n4	0.9952682	0.9961695	0.9953808	0.9951555	0.9952682	0.996	0.9974088	0.9977467	0.9968454

logg_g10_CFO.txt

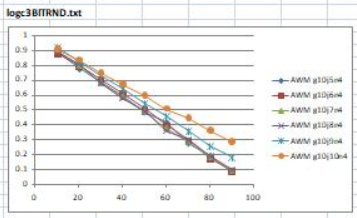
	2	3	4	5	6	7	8	9	10
AWM g50j6n4	1	1	1	1	0.9971671	1	1	1	1
AWM g20j6n4	0.9988385	1	0.9988385	0.9976772	0.9965157	0.999	0.9988385	0.9976772	0.9976772
AWM g10j6n4	0.9982886	0.9994295	0.9977182	0.9971477	0.9965773	0.999	0.9982886	0.9988591	0.9977182
AWM g8j6n4	0.9991232	0.9986848	0.997808	0.995616	0.997808	0.999	0.9982464	0.9995616	0.997808
AWM g5j6n4	0.9985632	0.9994253	0.9985632	0.9979885	0.9965518	0.999	0.9985632	0.9988506	0.9982759
AWM g4j6n4	0.9988921	0.9986705	0.9984489	0.997341	0.9977842	0.998	0.9986705	0.9995568	0.9982274
AWM g3j6n4	0.9979289	0.9989644	0.9982741	0.9972385	0.9977563	0.998	0.9989644	0.9996548	0.9982741
AWM g2j6n4	0.9984227	0.9984227	0.9979721	0.9970707	0.9975214	0.998	0.9985354	0.9995493	0.9980847

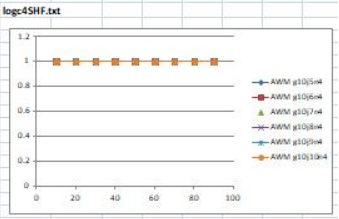
logg_g20_CFO.txt

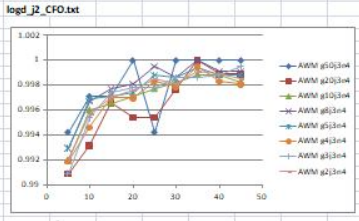
	2	3	4	5	6	7	8	9	10
AWM g50j6n4	1	1	1	1	1	1	1	1	1
AWM g20j6n4	0.9988385	1	1	0.9976772	0.9988385	0.999	1	0.9976772	0.9976772
AWM g10j6n4	0.9994295	1	1	0.9988591	0.9994295	0.999	0.9994295	0.9988591	0.9982886
AWM g8j6n4	0.9995616	0.9986848	0.9995616	0.9986848	0.9986848	1	1	0.9995616	1
AWM g5j6n4	0.9994253	1	0.9997127	0.9991379	0.9988506	1	0.9994253	0.9991379	0.9988506
AWM g4j6n4	0.9995568	0.9991137	0.9997784	0.9988921	0.9991137	0.999	1	0.9995568	0.9995568
AWM g3j6n4	0.9993097	0.9998274	0.9993097	0.9986193	0.999137	0.999	0.9996548	0.9996548	0.9993097
AWM g2j6n4	0.9992114	0.9990987	0.9990987	0.9985354	0.9990987	0.999	0.9995493	0.999662	0.9992114

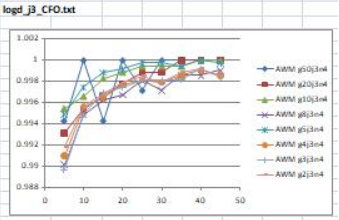
logg_g50_CFO.txt

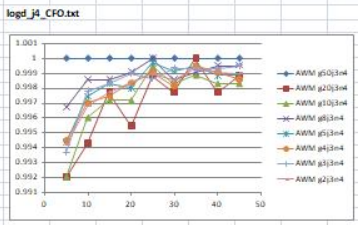
	2	3	4	5	6	7	8	9	10
AWM g50j6n4	1	1	1	1	1	1	1	1	1
AWM g20j6n4	0.9988385	1	1	0.9988385	1	1	1	1	0.9988385
AWM g10j6n4	0.9994295	1	1	0.9994295	1	1	1	1	0.9994295
AWM g8j6n4	1	1	0.9995616	0.9995616	1	1	1	1	0.9995616
AWM g5j6n4	0.9994253	1	0.9997127	0.9994253	1	1	1	0.9997127	0.9994253
AWM g4j6n4	0.9997784	1	0.9997784	0.9995568	0.9997784	1	1	1	0.9993352
AWM g3j6n4	1	1	0.9998274	0.9994822	0.9998274	0.999	0.9996548	1	0.9998274
AWM g2j6n4	0.9997747	1	0.9997747	0.9994367	0.999662	1	0.9997747	0.9998874	0.999324

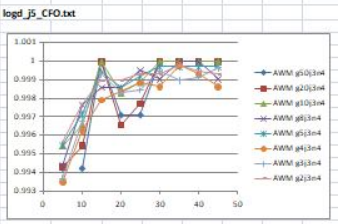


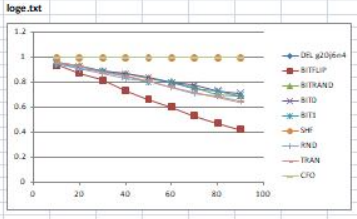


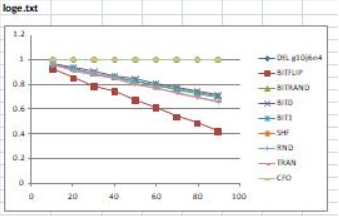


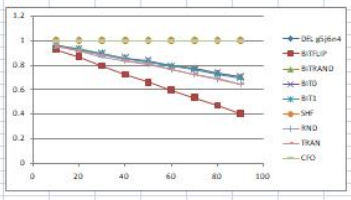


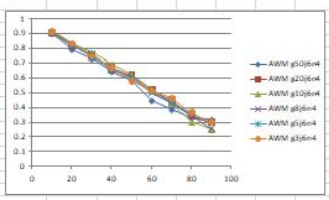


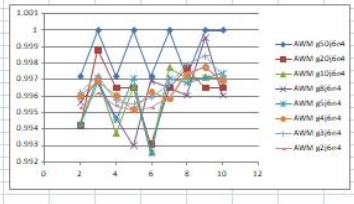


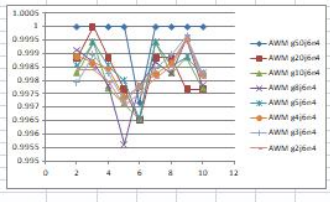


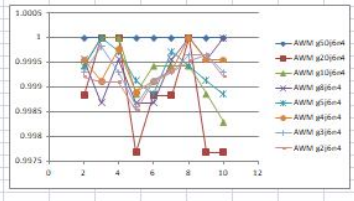


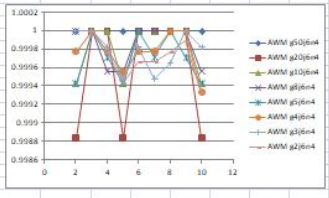












7.8 Εφαρμογή σεναρίων στο dataset 2: «Τιμές κρυπτονομισμάτων»

loga1DEL.txt

	10	20	30	40	50	60	70	80	90
AWM g50j20n4	1	1	1	1	1	1	1	1	1
AWM g20j20n4	1	1	1	1	1	1	1	1	1
AWM g10j20n4	1	1	1	1	1	1	1	1	1
AWM g8j20n4	1	1	1	1	1	1	1	1	1
AWM g5j20n4	1	1	1	1	1	1	1	1	1
AWM g3j20n4	1	1	1	1	1	1	1	1	1
AWM g2j20n4	1	1	1	1	1	1	1	1	1

loga2BITFLIP.txt

	10	20	30	40	50	60	70	80	90
AWM g50j20n4	0.94259393	0.88377035	0.83644225	0.7377746	0.7044649	0.6456414	0.5903614	0.5017718	0.45074415
AWM g20j20n4	0.93962157	0.88091266	0.8135782	0.75514746	0.6903172	0.6368948	0.58569837	0.5233723	0.46744573
AWM g10j20n4	0.9360078	0.87729037	0.81662965	0.7524986	0.7019711	0.6424209	0.5744031	0.52151585	0.46238202
AWM g8j20n4	0.9320035	0.8699581	0.8100066	0.75854087	0.6898832	0.6307031	0.5733965	0.5122328	0.4556976
AWM g5j20n4	0.9353261	0.87255436	0.81841034	0.7576766	0.6971467	0.63125	0.5779212	0.5172554	0.4552989
AWM g3j20n4	0.93351674	0.86974376	0.8145532	0.75373685	0.6927973	0.62886006	0.56993264	0.5097733	0.44990143
AWM g2j20n4	0.9337555	0.87395376	0.8162445	0.7514868	0.6965308	0.6328469	0.57252204	0.51420707	0.455837

loga3BITRAND.txt

	10	20	30	40	50	60	70	80	90
AWM g50j20n4	0.967399	0.9431765	0.9064493	0.8965273	0.8511694	0.8093551	0.80155915	0.7590361	0.7278526
AWM g20j20n4	0.9671675	0.9376739	0.9006678	0.88146913	0.84696716	0.8174736	0.79660547	0.7726767	0.7273233
AWM g10j20n4	0.9693226	0.9378123	0.9119933	0.8759023	0.8460578	0.82398865	0.79386455	0.7626319	0.7323709
AWM g8j20n4	0.97024465	0.937624	0.90379107	0.8784439	0.8410844	0.818272	0.7797002	0.7551245	0.7242671
AWM g5j20n4	0.964606	0.9356658	0.9109375	0.88158965	0.8480299	0.8236413	0.79110056	0.7605299	0.727038
AWM g3j20n4	0.96612185	0.9361449	0.9053055	0.875657	0.8429287	0.81841326	0.7888469	0.7580486	0.72330815
AWM g2j20n4	0.9669879	0.9373073	0.9065253	0.87888217	0.8496421	0.8181167	0.7819383	0.7588381	0.7288546

loga4BIT0.txt

	10	20	30	40	50	60	70	80	90
AWM g50j20n4	0.9744862	0.9418852	0.9007796	0.8640884	0.8534562	0.8176796	0.80829495	0.77235025	0.73038673
AWM g20j20n4	0.96688926	0.9325125	0.9073456	0.8828089	0.84074074	0.81851345	0.78962964	0.751482	0.7273142
AWM g10j20n4	0.9632149	0.9323902	0.90191656	0.8786785	0.84802705	0.8214205	0.7889515	0.7609921	0.7287486
AWM g8j20n4	0.96763813	0.9361913	0.90756303	0.87234044	0.8451636	0.8154836	0.7807974	0.7530842	0.71768284
AWM g5j20n4	0.9669837	0.9393342	0.90638155	0.8770219	0.8487702	0.8148616	0.78639483	0.7586971	0.7265677
AWM g3j20n4	0.96649146	0.93367964	0.9062893	0.8747125	0.8438807	0.81459993	0.7858587	0.75279987	0.7184307
AWM g2j20n4	0.9674559	0.9388897	0.9063051	0.8777673	0.8498361	0.8165162	0.7877767	0.7580044	0.7276932

loga5BIT1.txt

	10	20	30	40	50	60	70	80	90
AWM g50j20n4	0.962438	0.93267184	0.9210431	0.8929837	0.864635	0.83715095	0.7930546	0.76399714	0.7328136
AWM g20j20n4	0.9727323	0.9443517	0.9092933	0.8903729	0.8558709	0.82248193	0.7982749	0.75792986	0.7359488
AWM g10j20n4	0.9705719	0.94392	0.912965	0.87895614	0.8480031	0.8260689	0.7973348	0.7679067	0.740422
AWM g8j20n4	0.9667181	0.94192195	0.90379107	0.8784439	0.84626406	0.8146352	0.78410846	0.7579899	0.72845495
AWM g5j20n4	0.96868205	0.93974185	0.9068614	0.8776495	0.85013586	0.8184783	0.7879076	0.76154894	0.7298913
AWM g3j20n4	0.9660398	0.9384034	0.9096994	0.87820303	0.84703517	0.82017905	0.7910644	0.7598344	0.7300427
AWM g2j20n4	0.96596915	0.93876654	0.9093134	0.8751927	0.8486784	0.8182269	0.7865914	0.7582048	0.7253304

loga6SHF.txt

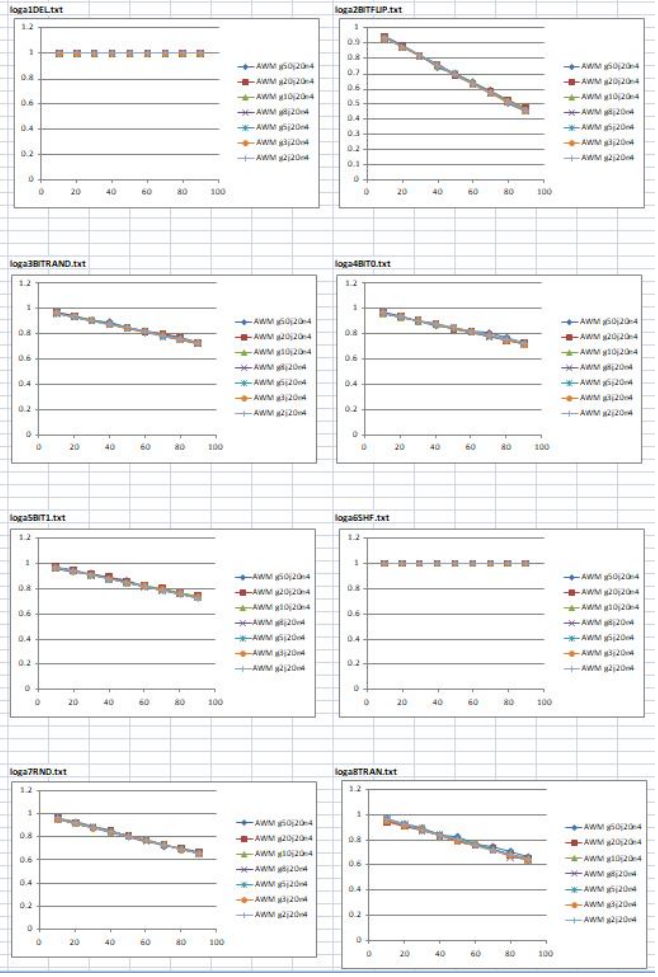
	10	20	30	40	50	60	70	80	90
AWM g50j20n4	1	1	1	1	1	1	1	1	1
AWM g20j20n4	1	1	1	1	1	1	1	1	1
AWM g10j20n4	1	1	1	1	1	1	1	1	1
AWM g8j20n4	1	1	1	1	1	1	1	1	1
AWM g5j20n4	1	1	1	1	1	1	1	1	1
AWM g3j20n4	1	1	1	1	1	1	1	1	1
AWM g2j20n4	1	1	1	1	1	1	1	1	1

loga7RND.txt

	10	20	30	40	50	60	70	80	90
AWM g50j20n4	0.95393336	0.929837	0.8858965	0.8561304	0.8029766	0.77888024	0.71793056	0.69808644	0.6676116
AWM g20j20n4	0.9618809	0.9229271	0.8795214	0.84891486	0.81218696	0.7673901	0.73761827	0.69616026	0.663606
AWM g10j20n4	0.9632149	0.92379236	0.8831205	0.84716827	0.8145475	0.77262634	0.7322321	0.7023876	0.6624098
AWM g8j20n4	0.9599957	0.9205422	0.8824113	0.8530968	0.8116597	0.76603484	0.7307692	0.6958343	0.6573727
AWM g5j20n4	0.9587636	0.92228264	0.88410324	0.8444294	0.80788046	0.7782609	0.7410326	0.70203805	0.6648777
AWM g3j20n4	0.95741624	0.92029405	0.8821452	0.8471584	0.8083314	0.77414584	0.7324655	0.69431135	0.65744084
AWM g2j20n4	0.9609581	0.9202092	0.88221365	0.84636563	0.80944383	0.7753579	0.73177314	0.6943557	0.6585903

loga8TRAN.txt

	10	20	30	40	50	60	70	80	90
AWM g50j20n4	0.97109824	0.92774564	0.8959538	0.8410405	0.8265896	0.76878613	0.7485549	0.7109827	0.6676301
AWM g20j20n4	0.94736844	0.91075516	0.8764302	0.83524024	0.79748285	0.7562929	0.7276888	0.6773455	0.64416474
AWM g10j20n4	0.9564719	0.9203895	0.8825888	0.83390605	0.8087056	0.7726231	0.731386	0.69072163	0.6575029
AWM g8j20n4	0.9600376	0.9186648	0.8735308	0.83215797	0.79078513	0.75787497	0.7240245	0.66572636	0.64221907
AWM g5j20n4	0.9623869	0.92335975	0.88687783	0.8410633	0.8034502	0.7630091	0.7231335	0.68636876	0.64621043



logb1DEL.txt

	10	12	14	16	18	20
AWM g50j10n4	1	1	1	1	1	1
AWM g20j10n4	1	1	1	1	1	1
AWM g10j10n4	1	1	1	1	1	1
AWM g8j10n4	1	1	1	1	1	1
AWM g5j10n4	1	1	1	1	1	1
AWM g3j10n4	1	1	1	1	1	1

logb2BITFLIP.txt

	10	12	14	16	18	20
AWM g50j10n4	0.6881644	0.686747	0.71216954	0.70233876	0.7526577	0.7746279
AWM g20j10n4	0.67668337	0.67723984	0.6917084	0.72036725	0.7520868	0.75709516
AWM g10j10n4	0.6957246	0.69891727	0.69752914	0.70960575	0.7351471	0.77082175
AWM g8j10n4	0.6949526	0.695173	0.69175667	0.7042098	0.7420101	0.7676879
AWM g5j10n4	0.6929348	0.684375	0.6910326	0.70774454	0.741644	0.7675348
AWM g3j10n4	0.6963699	0.693249	0.68951213	0.7118101	0.7403088	0.76852006

logb3BITRND.txt

	10	12	14	16	18	20
AWM g50j10n4	0.8476258	0.8299079	0.85329556	0.86534375	0.8724309	0.892275
AWM g20j10n4	0.84696716	0.8350028	0.845576	0.8519755	0.8625487	0.8875904
AWM g10j10n4	0.8502221	0.83897835	0.83772904	0.8571627	0.8649362	0.88992226
AWM g8j10n4	0.84626406	0.8435089	0.85078245	0.85948867	0.87436634	0.8759092
AWM g5j10n4	0.8435462	0.8493886	0.84701085	0.85944295	0.867731	0.88423914
AWM g3j10n4	0.8423538	0.8460496	0.8438732	0.8529895	0.8609491	0.88538927

logb4SHF.txt

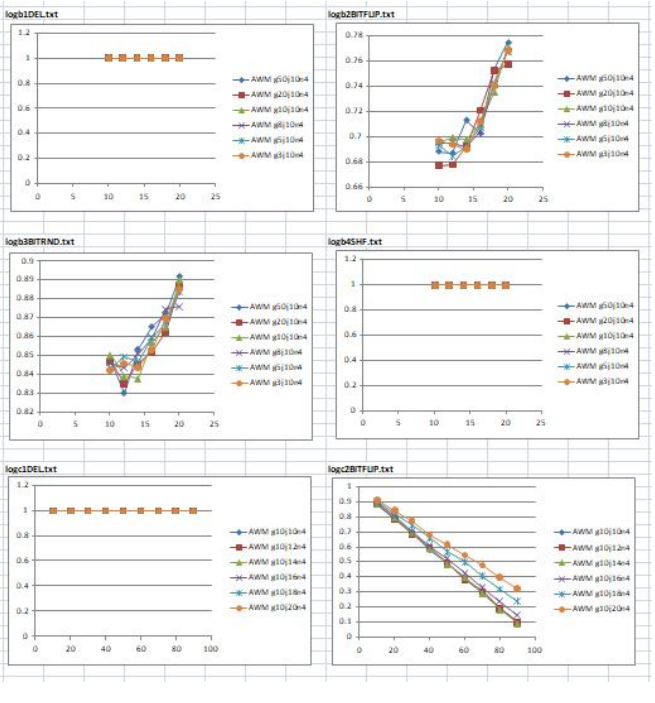
	10	12	14	16	18	20
AWM g50j10n4	1	1	1	1	1	1
AWM g20j10n4	1	1	1	1	1	1
AWM g10j10n4	1	1	1	1	1	1
AWM g8j10n4	1	1	1	1	1	1
AWM g5j10n4	1	1	1	1	1	1
AWM g3j10n4	1	1	1	1	1	1

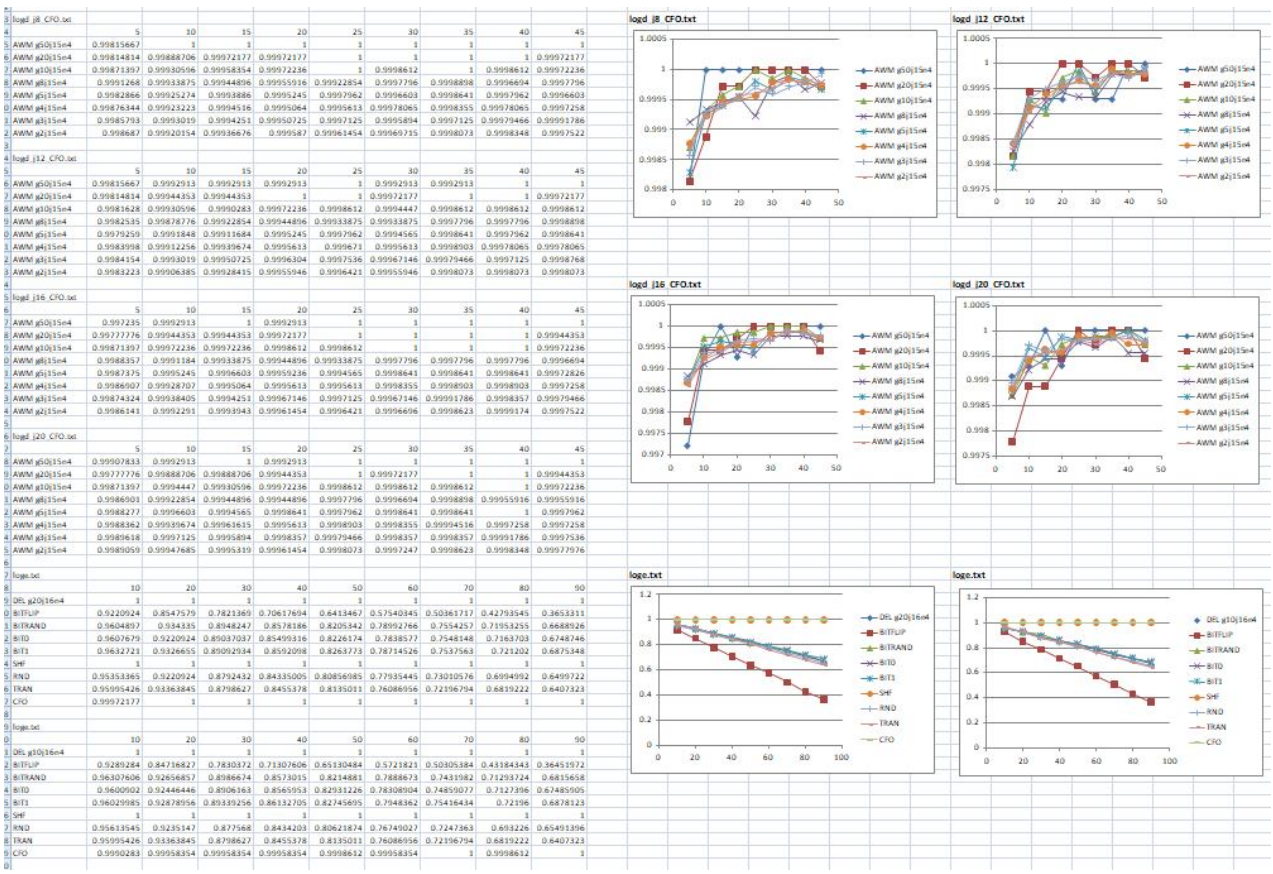
logc1DEL.txt

	10	20	30	40	50	60	70	80	90
AWM g10j10n4	1	1	1	1	1	1	1	1	1
AWM g10j12n4	1	1	1	1	1	1	1	1	1
AWM g10j14n4	1	1	1	1	1	1	1	1	1
AWM g10j16n4	1	1	1	1	1	1	1	1	1
AWM g10j18n4	1	1	1	1	1	1	1	1	1
AWM g10j20n4	1	1	1	1	1	1	1	1	1

logc2BITFLIP.txt

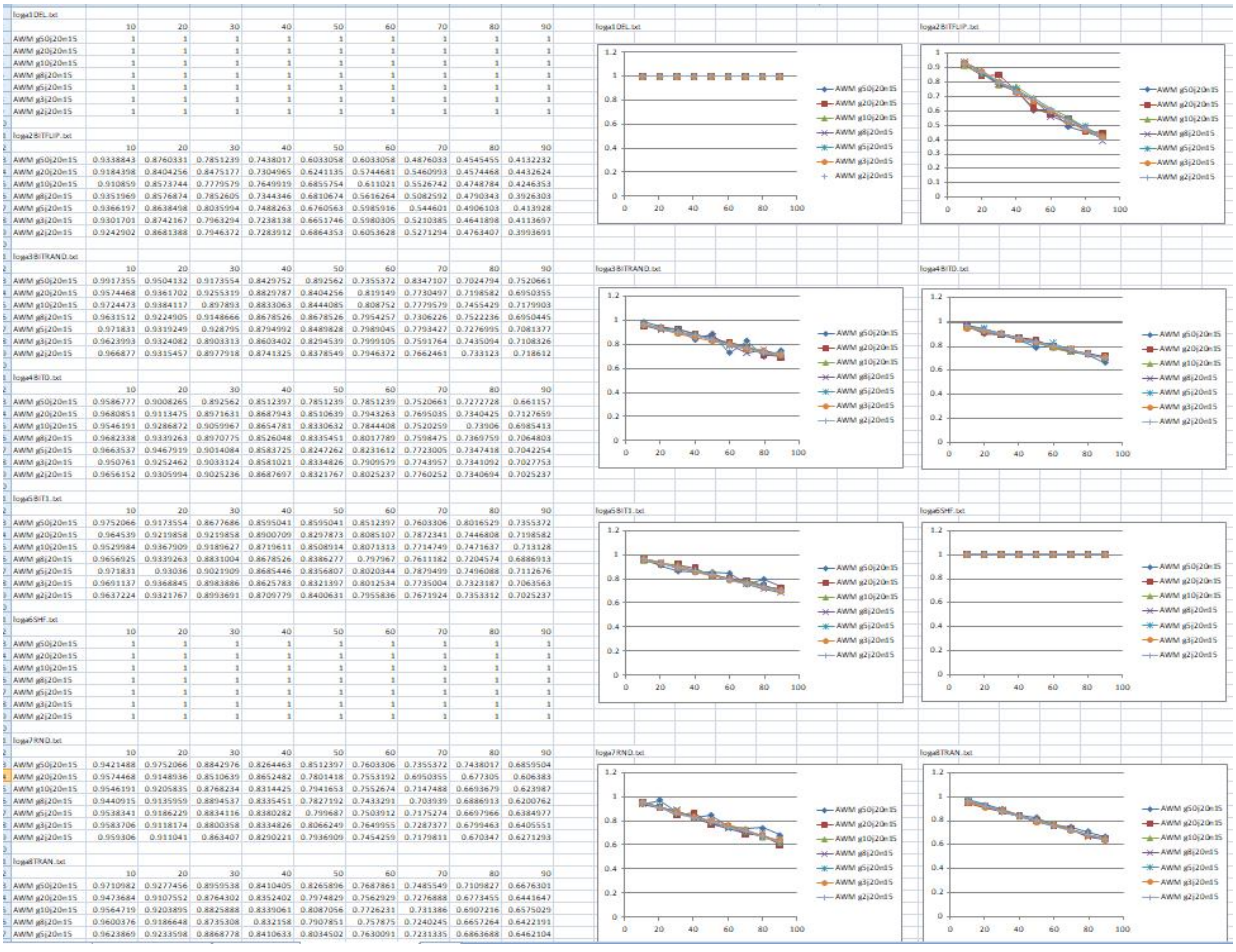
	10	20	30	40	50	60	70	80	90
AWM g10j10n4	0.88249797	0.78834024	0.69323474	0.57984114	0.49055059	0.38824978	0.2885511	0.18816763	0.09737058
AWM g10j12n4	0.8887976	0.7869077	0.6877568	0.5946316	0.49452204	0.3812654	0.3034785	0.19241303	0.09682279
AWM g10j14n4	0.89003015	0.79745275	0.69857574	0.5927143	0.48562038	0.39961654	0.2917009	0.18255273	0.08942755
AWM g10j16n4	0.89824706	0.80786085	0.70651877	0.6084634	0.52423996	0.42495206	0.33100522	0.2363736	0.14393316
AWM g10j18n4	0.906327	0.82203356	0.74437967	0.6615996	0.5684744	0.5021912	0.40892905	0.3203232	0.23870173
AWM g10j20n4	0.91536584	0.84401536	0.7771843	0.68186796	0.6185976	0.5468365	0.48110107	0.3978362	0.32224047

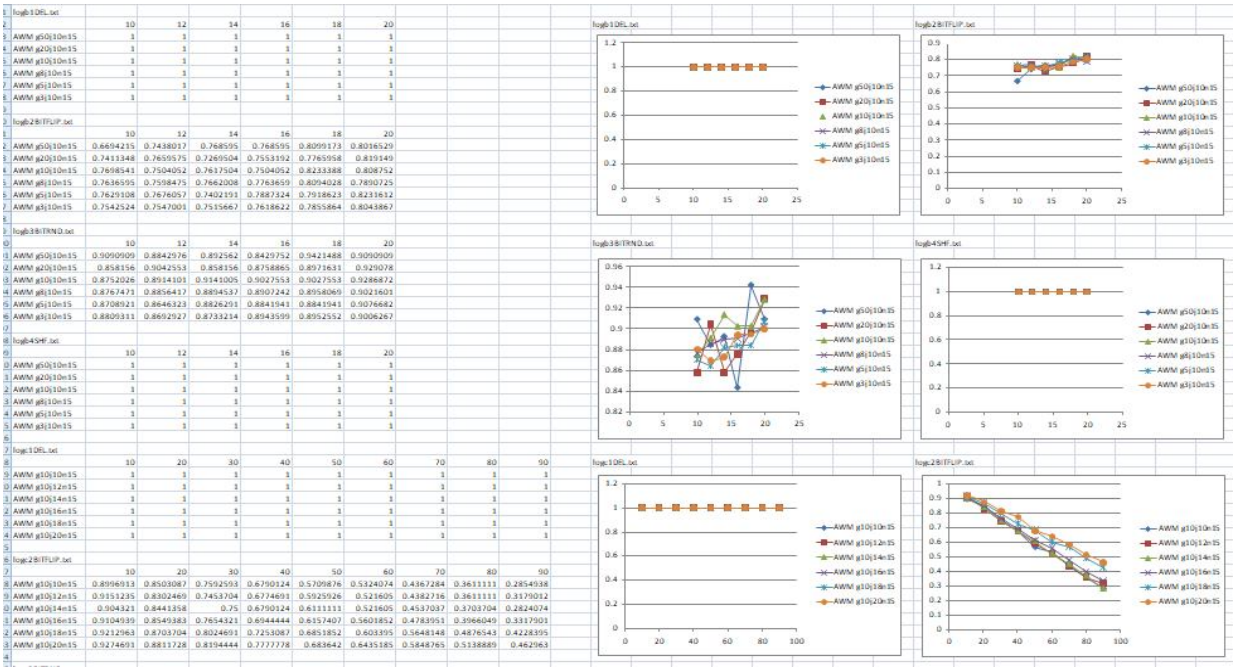


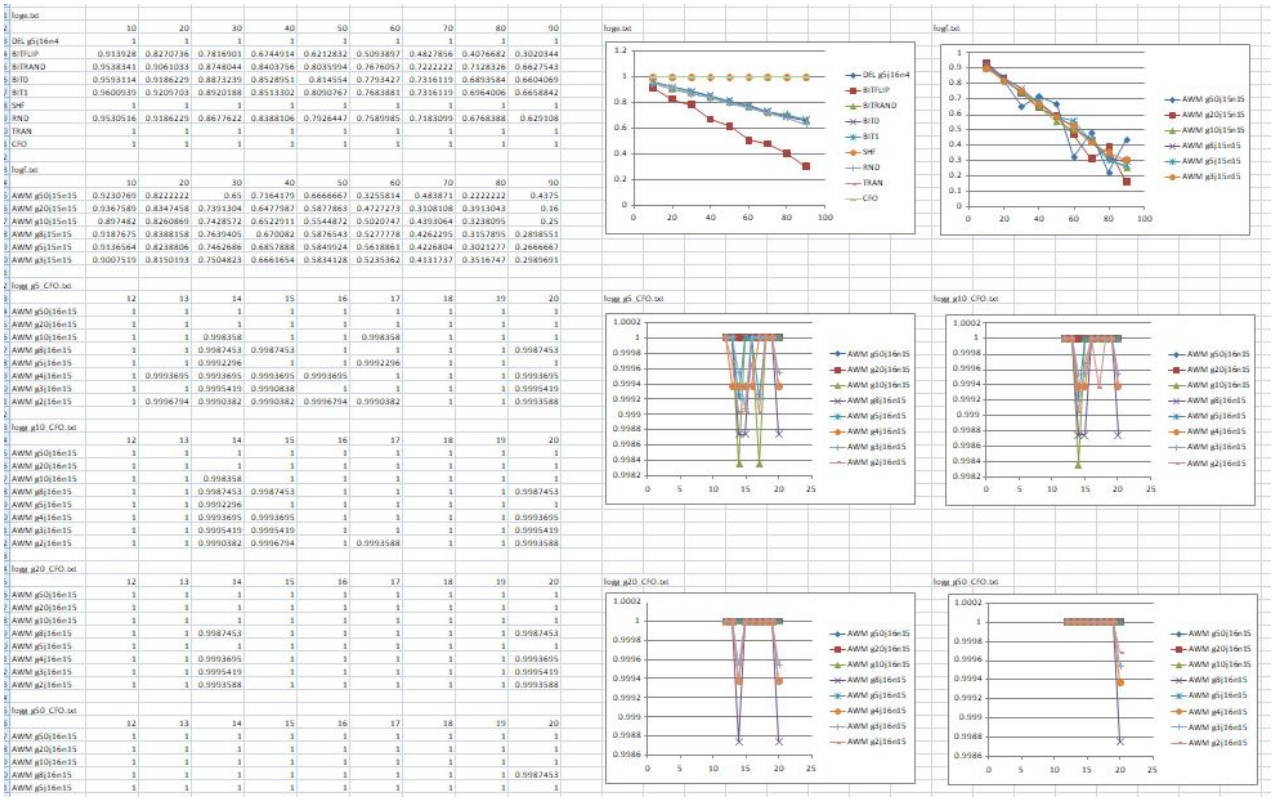

logd_j8_CFO.txt
logd_j12_CFO.txt
logd_j16_CFO.txt
logd_j20_CFO.txt
loge.txt
loge.txt


loge.txt
logf.txt
logg_g5_CFO.txt
logg_g10_CFO.txt
logg_g20_CFO.txt
logg_g50_CFO.txt

7.9 Εφαρμογή σεναρίων στο dataset 3: «Τιμές χρηματιστηρίου»







Παράρτημα ΙΙ : Κρυπτογραφία – Στεγανογραφία - Υδατογραφία

Αυτό το παράρτημα παρουσιάζει ιστορικές πληροφορίες για την κρυπτογραφία και την στεγανογραφία που αποτελούν τη βάση για την υδατογραφία. Στη στεγανογραφία, το μήνυμα που αποστέλλεται είναι κρυμμένο ενώ στην κρυπτογραφία μετασχηματίζεται με τη χρήση ενός μυστικού κλειδιού, ώστε να είναι ακατανόητο σε οποιοδήποτε ξένο. Οι τρεις έννοιες έχουν ως δεύτερο συστατικό τη λέξη «γραφή» που σχετίζεται με το «μήνυμα» ενώ εξυπηρετούν το στόχο της εμπιστευτικότητας με διαφορετικό τρόπο. Η Εικόνα 34 αποτυπώνει τις 3 έννοιες.

Εικόνα 34: Κρυπτο vs Στεγανο vs Υδατο : ΓΡΑΦΗΣΗ

ΚΡΥΠΤΟΓΡΑΦΙΑ

Η λέξη κρυπτογραφία προέρχεται από τα συνθετικά «κρυπτός» και «γραφή» και ασχολείται με τη μελέτη, την ανάπτυξη και τη χρήση τεχνικών κωδικοποίησης και αποκωδικοποίησης του περιεχομένου εμπιστευτικών μηνυμάτων τα οποία ανταλλάσσονται με τρόπο ώστε η κατανόηση του περιεχομένου των μηνυμάτων να είναι δυνατή μόνο από τους αυτούς που συμμετέχουν στην επικοινωνία, δηλαδή τον αποστολέα και τον παραλήπτη [83]

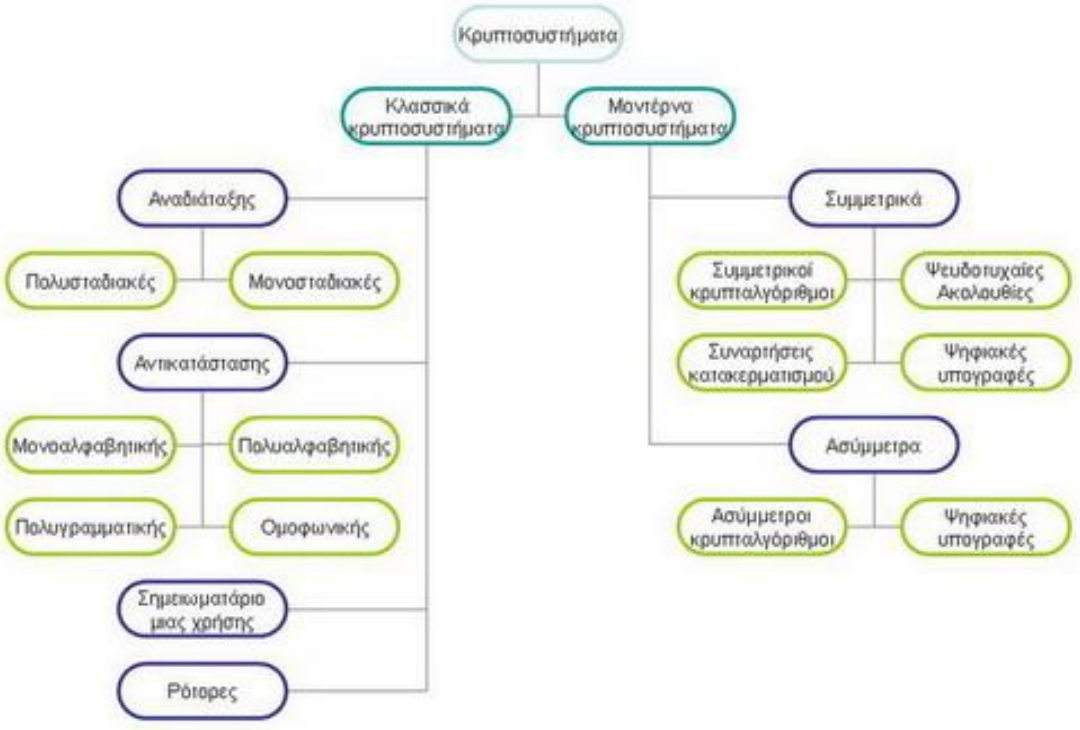
Η ανάπτυξη της κρυπτογραφίας χωρίζεται σε 3 περιόδους.

Στην πρώτη περίοδο (1900π.Χ. – 1900μ.Χ.) αναπτύχθηκαν πρωτότυπες για την εποχή μέθοδοι κρυπτογράφησης που βασίστηκαν σε αντικαταστάσεις γραμμάτων και στηρίχθηκαν στην ευρηματικότητα των δημιουργών τους. Η πρώτη στρατιωτική εφαρμογή της κρυπτογραφίας έγινε τον $5^{ο}$ π.Χ. αιώνα από τους Σπαρτιάτες. Αυτοί εφηύραν την «σκυτάλη» στην οποία χρησιμοποίησαν τη μέθοδο της μετάθεσης τυλίγοντας λωρίδα με το μήνυμα γύρω από τη σκυτάλη. Η περίμετρος έπαιζε το ρόλο του κλειδιού και καθόριζε αν μπορεί να

αποκρυπτογραφηθεί το μήνυμα. Με την ίδια μέθοδο, ο Ιούλιος Καίσαρας έγραφε στον Κικέρωνα και σε άλλους φίλους του, αντικαθιστώντας τα γράμματα του κειμένου, με γράμματα, που βρίσκονται 3 θέσεις μετά, στο Λατινικό Αλφάβητο. Στη διάρκεια του Μεσαίωνα, η κρυπτολογία απαγορεύτηκε αφού συνδέθηκε με τον αποκρυφισμό και τη μαύρη μαγεία και αυτό καθυστέρηση της ανάπτυξης της. Οι Άραβες τον 14^{ο} αιώνα επινόησαν τις πρώτες μεθόδους κρυπτανάλυσης χρησιμοποιώντας τις συχνότητες εμφάνισης γραμμάτων σε κείμενα. Η κρυπτογραφία σημείωσε ανάπτυξη στους επόμενους αιώνες λόγω των πολέμων. Ο Γάλλος Vigenere δημιούργησε τον πίνακα πολυαλφαβητικής αντικατάστασης ως εξέλιξη της αρχικής μετάθεσης.

Στη δεύτερη περίοδο(1900-1950 μ.Χ) η κρυπτογραφία εξελίχθηκε με γρήγορους ρυθμούς εν' μέσω των δύο παγκόσμιων πολέμων που έθεσαν ως ανάγκη την ασφαλή μετάδοση ζωτικών πληροφοριών στρατιωτικού περιεχομένου. Το γερμανικό σύστημα Enigma και η αντίστοιχη ηλεκτρομηχανική κρυπτογραφική μηχανή του, δοκιμάσθηκαν στο πεδίο ενώ στην πορεία παραβιάστηκε μετά από επιθέσεις και πειράματα.

Στην τρίτη περίοδο (1950 - σήμερα), η ανάπτυξη των μαθηματικών, της ηλεκτρονικής και των υπολογιστικών συστημάτων συντέλεσε στην υλοποίηση συμμετρικών και ασύμμετρων αλγορίθμων κρυπτογραφίας, συναρτήσεων κατακερματισμού, τεχνολογιών blockchain με εφαρμογές στις επικοινωνίες και σε πολλούς κλάδους της οικονομίας. Η κρυπτογραφία ως διεπιστημονικός τομέας έχει προοπτικές επέκτασης σε χώρους όπως η κβαντική μηχανική. Τα κρυπτοσυστήματα που υπάρχουν σήμερα είναι πολλά και φαίνονται ενδεικτικά στην Εικόνα 35.



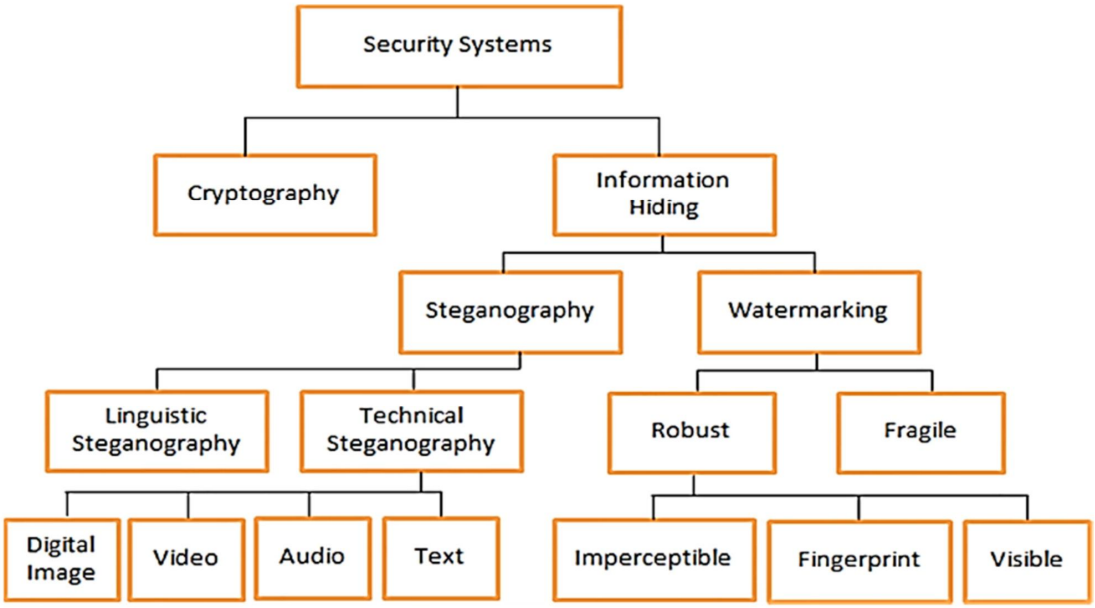
Εικόνα 35: Κατηγορίες κρυπτογραφίας

ΣΤΕΓΑΝΟΓΡΑΦΙΑ

Η στεγανογραφία [84] προέρχεται από τις λέξεις «στεγανός» και «γραφή» και ως διαδικασία αποκρύπτει ένα μήνυμα μέσα σε ένα διαφορετικό μέσο χωρίς να γίνεται γνωστή η

ύπαρξή του. Η στεγανογραφία ως ιδέα ξεκίνησε στην αρχαιότητα και αυτό έχει καταγραφεί σε ιστορικά γεγονότα. Ο Ηρόδοτος, αναφέρει ότι ο Δημάρατος για να ειδοποιήσει τη Σπάρτη ότι ο Ξέρξης προτίθετο να εισβάλει στην Ελλάδα, έγραψε το μήνυμά σε ξύλινη πινακίδα, αφού έξυσε το κερί που αυτή είχε και την οποία μετά κάλυψε πάλι με κερί. Οι πινακίδες φαίνονταν λευκές και αχρησιμοποίητες και με αυτό το τρόπο δεν κίνησαν υποψίες κατά τη μεταφορά τους.

Η στεγανογραφία σήμερα υλοποιείται με την τεχνολογία της πληροφορικής κρύβοντας ψηφιοποιημένες μορφές μηνυμάτων μέσα σε άλλες μορφές : κείμενο μέσα σε εικόνα, εικόνα μέσα σε σήμα, βιομετρικά στοιχεία μέσα σε βάση δεδομένων κ.α. Η στεγανογραφία προσφέρει προστασία του μηνύματος μέσω της μυστικότητας. Αν γίνει αντιληπτή η ύπαρξη του μηνύματος, είναι θέμα χρόνου η ανάκτησή του από κακόβουλους. Ο συνδυασμός στεγανογραφίας και κρυπτογραφίας οδηγεί σε ασφαλέστερες λύσεις όπου τα μηνύματα πρώτα κωδικοποιούνται και στη συνέχεια κρύβονται σε μια εικόνα ή σε ένα βίντεο. Οι τεχνικές της απόκρυψης πληροφορίας απουπώνονται στην Εικόνα 36.



Εικόνα 36: Κατηγορίες στεγανογραφίας

ΥΔΑΤΟΓΡΑΦΙΑ

Η υδατογραφία [85] προέρχεται από τις λέξεις «ύδωρ» και «γραφή». Το υδατογράφημα ή υδατόσημο ξεκίνησε ως εικόνα ή σχέδιο ενσωματωμένο σε χαρτί, που μπορούσε να αποκαλυφθεί με έντονο φως. Η κατασκευή του βασίστηκε στη διάταξη των ινών του χαρτιού, εισάγοντας διακυμάνσεις στην πυκνότητα. Τα υδατογραφήματα ξεκίνησαν τον 13ο αιώνα σε χαρτονομίσματα, διαβατήρια, γραμματόσημα, έγγραφα υψηλής αξίας και σε απλά χαρτιά αλληλογραφίας, ως μέσο αναγνώρισης του οίκου κατασκευής. Τα υδατόσημα στα παλαιά χειρόγραφα προσδιορίζουν τον κατασκευαστή του χαρτιού και την εποχή της κατασκευής του.

Τα φυσικά και τα ορατά υδατογραφήματα προσφέρουν ασφάλεια, τεκμήρια αυθεντικότητας και τα συναντάμε σε τίτλους σπουδών, σε εισιτήρια, σε συμβόλαια, σε εκτυπώσεις έργων τέχνης.

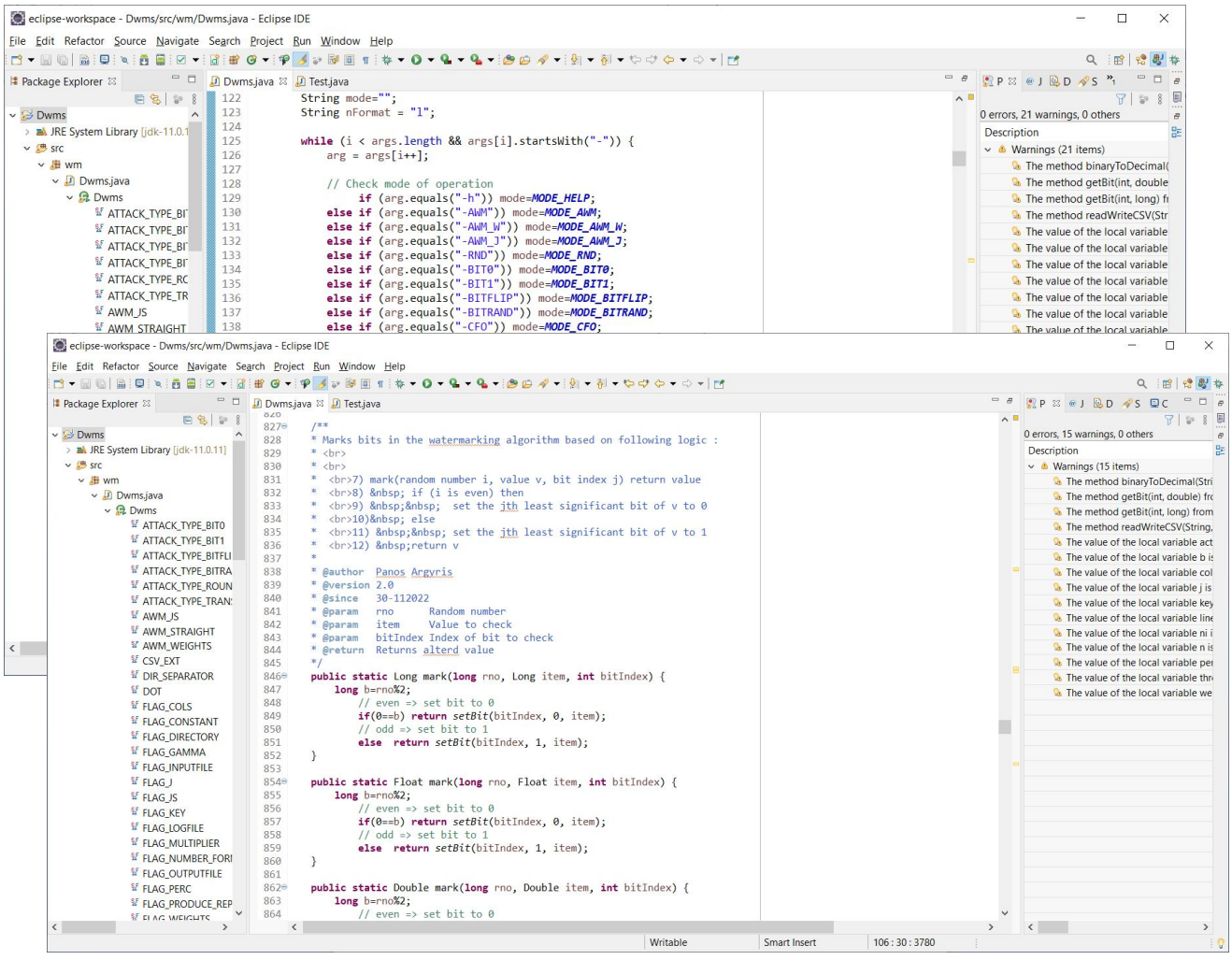
Το ψηφιακό υδατογράφημα αποτελεί σύγχρονη μέθοδος ψηφιακής κωδικοποίησης ηλεκτρονικών αρχείων που ενσωματώνει ένα κρυφό μήνυμα στο περιεχόμενο ενδιαφέροντος.

Δανείζεται τη λογική της στεγανογραφίας αλλά ενσωματώνει το μήνυμα στο ίδιο μέσο εξασφαλίζοντας την ανθεκτικότητά του, δηλαδή την αδυναμία αφαίρεσής του.

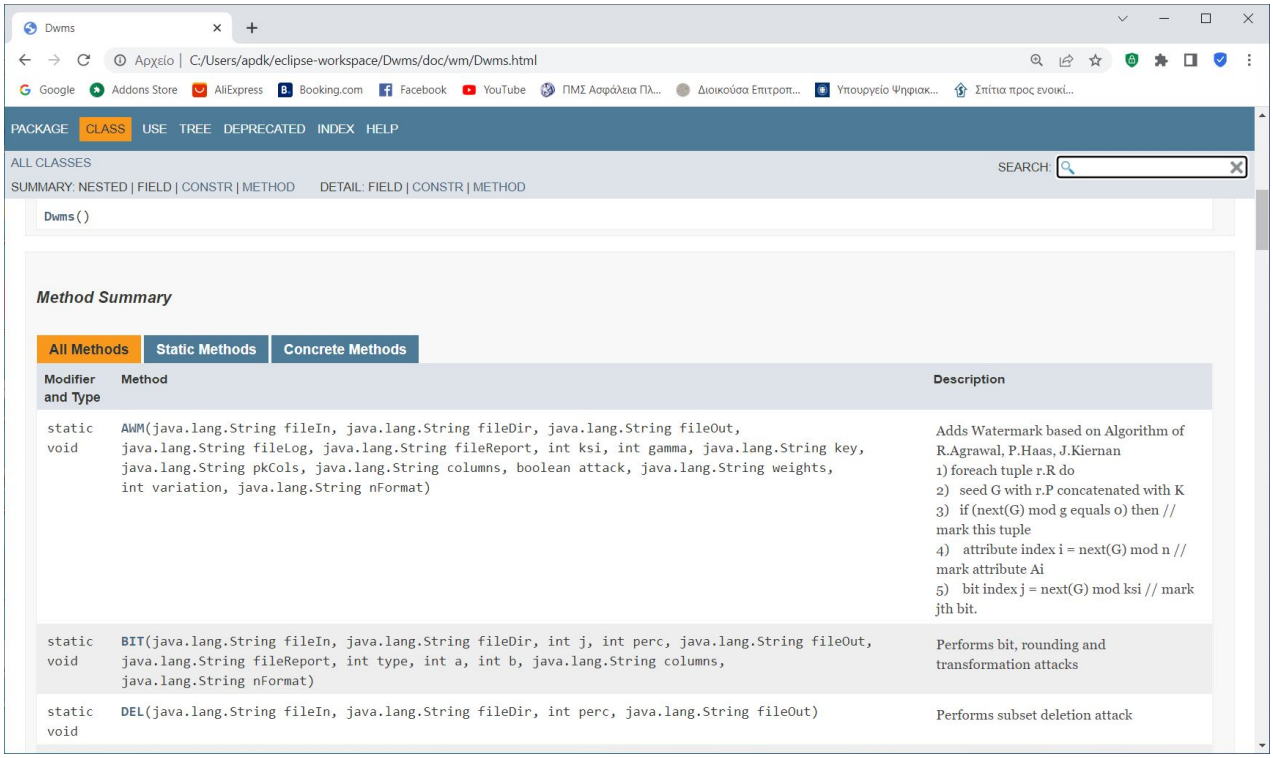
Ένα ψηφιακό αντίγραφο είναι ταυτόσημο με το πρωτότυπο ενώ η ψηφιακή υδατογράφηση έρχεται ως μορφή παθητικής προστασίας για να σημάνει ένα ψηφιακό έργο με τρόπο διακριτικό και να παρέχει δυνατότητες α) ανίχνευσης με στόχο την αναγνώριση του δημιουργού, β) ιχνηλάτισης της πηγής που αυτό προήλθε και τυχών παράνομων αντιγραφών (εφόσον έχει προστεθεί υδατογράφημα σε κάθε σημείο διανομής) και γ) επιβεβαίωσης ακεραιότητας μέσα από εύθραυστα σχήματα που είναι ευαίσθητα σε μεταβολές.

Παράρτημα ΙΙΙ : Τεκμήρια από το πρόγραμμα, τα datasets, τα πειράματα και τα αποτελέσματα

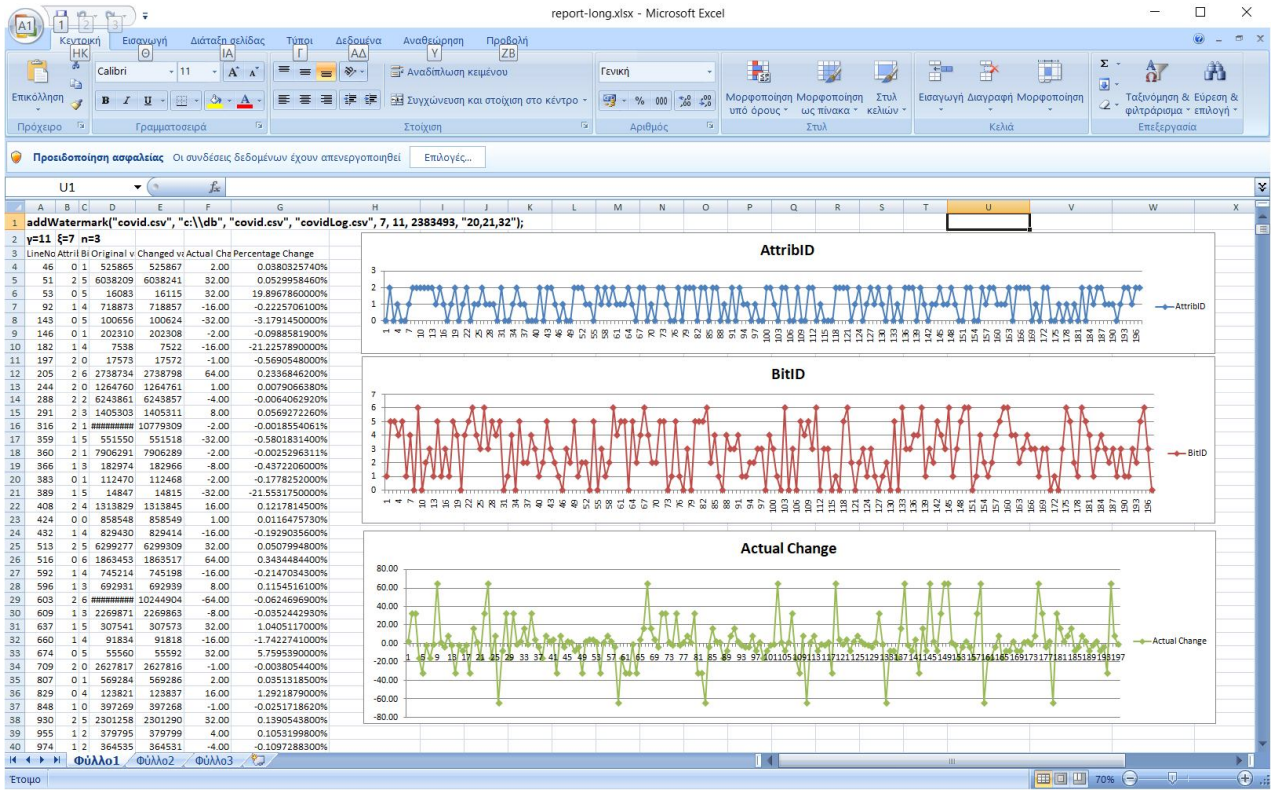
Αυτό το παράρτημα παρουσιάζει εικόνες από το εργαλείο ανάπτυξης Eclipse IDE v.2021-06(4.20.0), επίμαχα σημεία του κώδικα και τεκμηρίωση του προγράμματος. Επίσης παραθέτει κάποια παραδείγματα σημαντικών scripts στην αυτοματοποιημένη εκτέλεση των πειραμάτων.



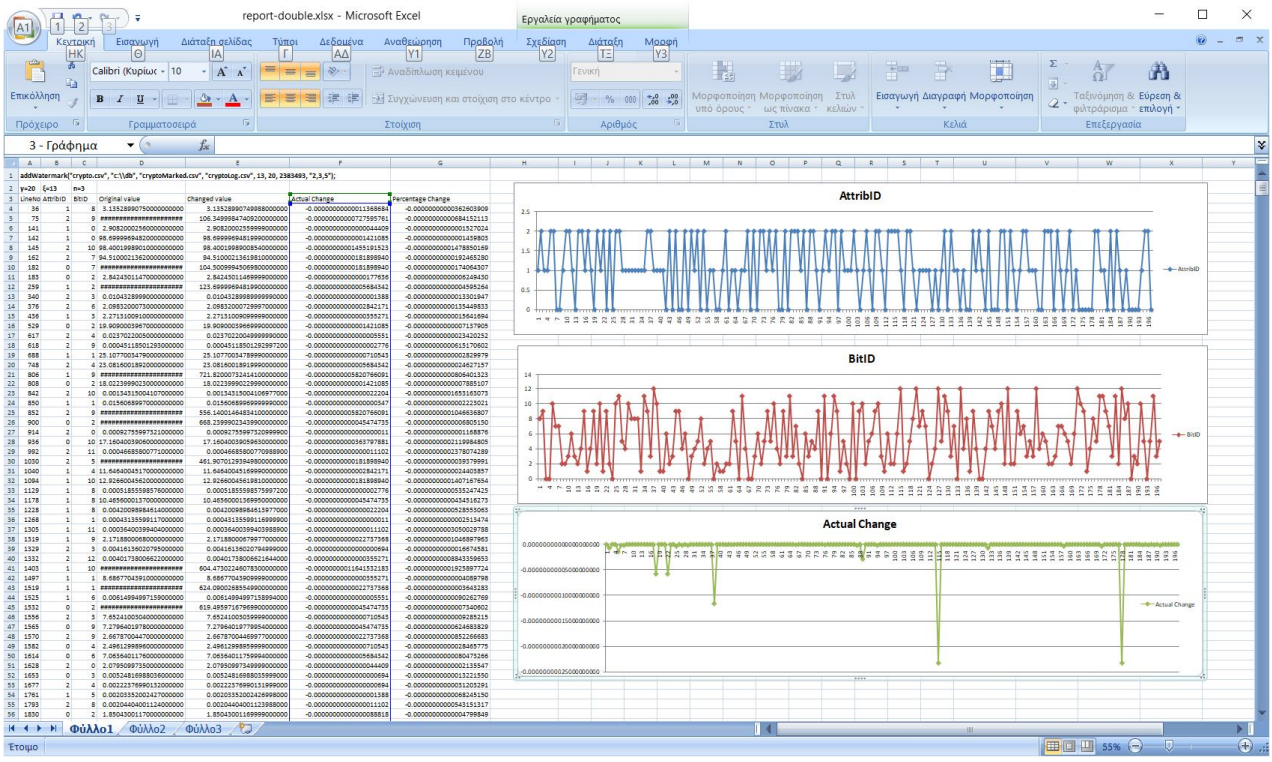
Εικόνα 37: Ανάπτυξη του λογισμικού με το Eclipse IDE



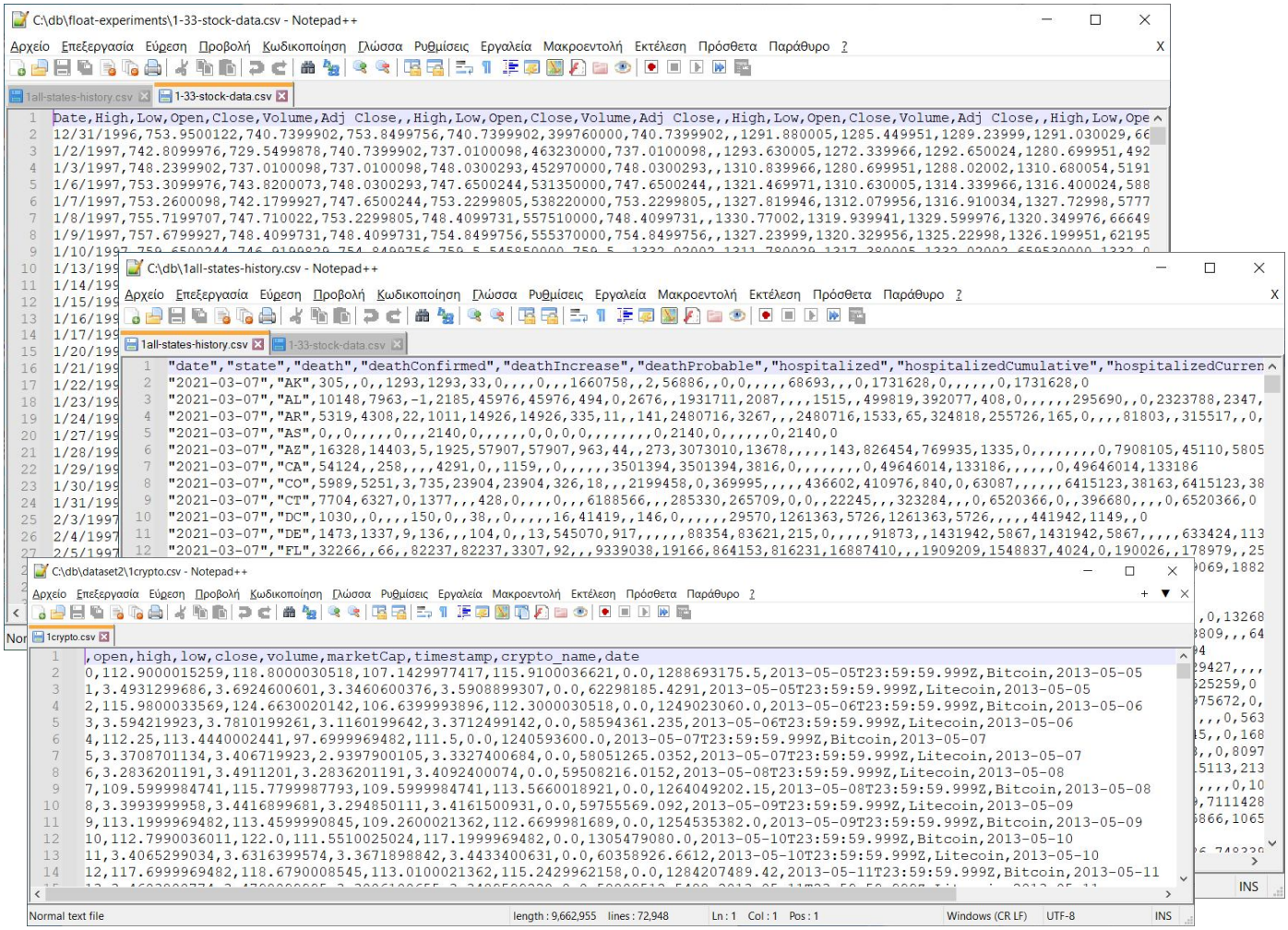
Εικόνα 38: Javadoc του λογισμικού



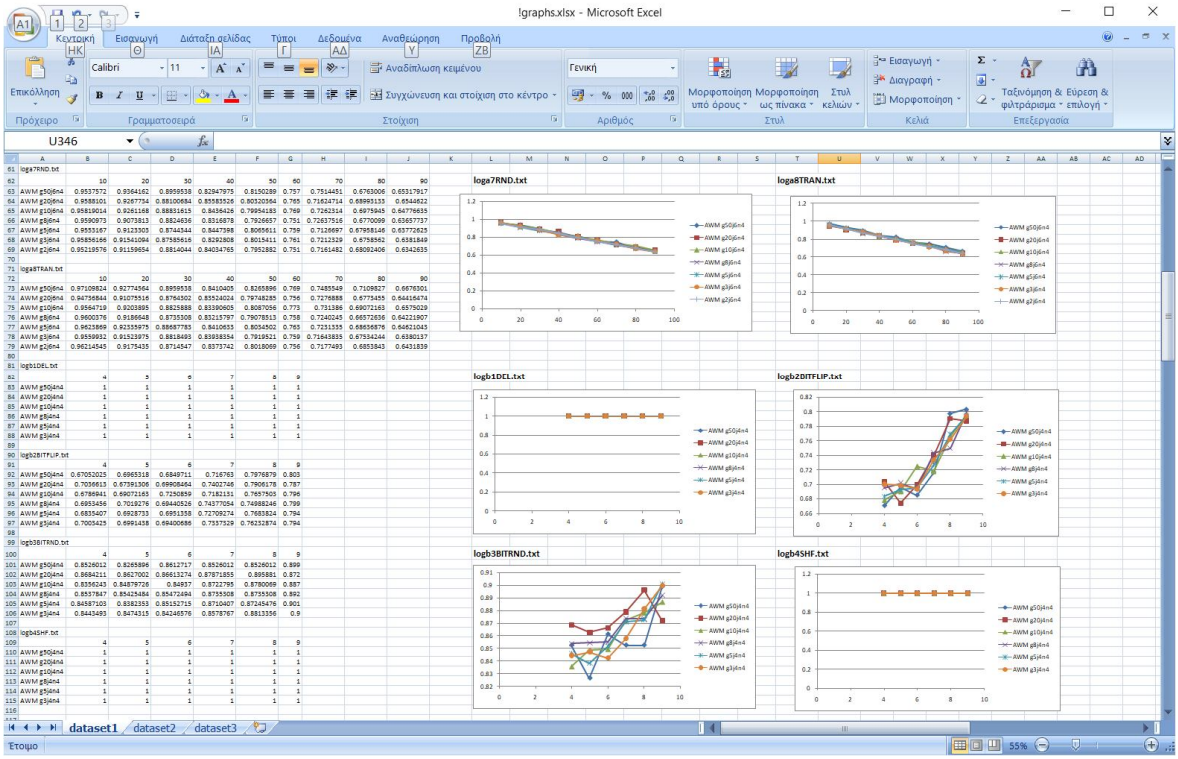
Εικόνα 39: Έλεγχοι του αλγόριθμου σε dataset ακεραίων αριθμών



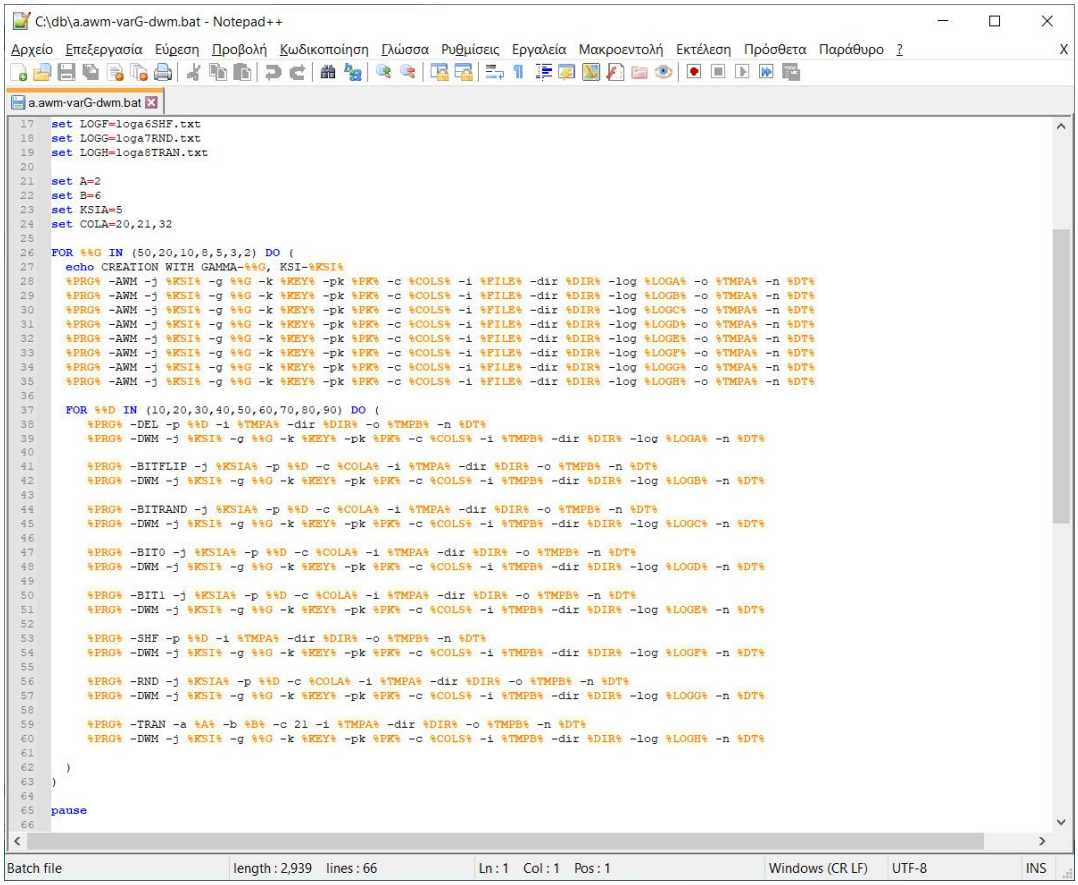
Εικόνα 40: Έλεγχοι του αλγόριθμου σε dataset δεκαδικών αριθμών



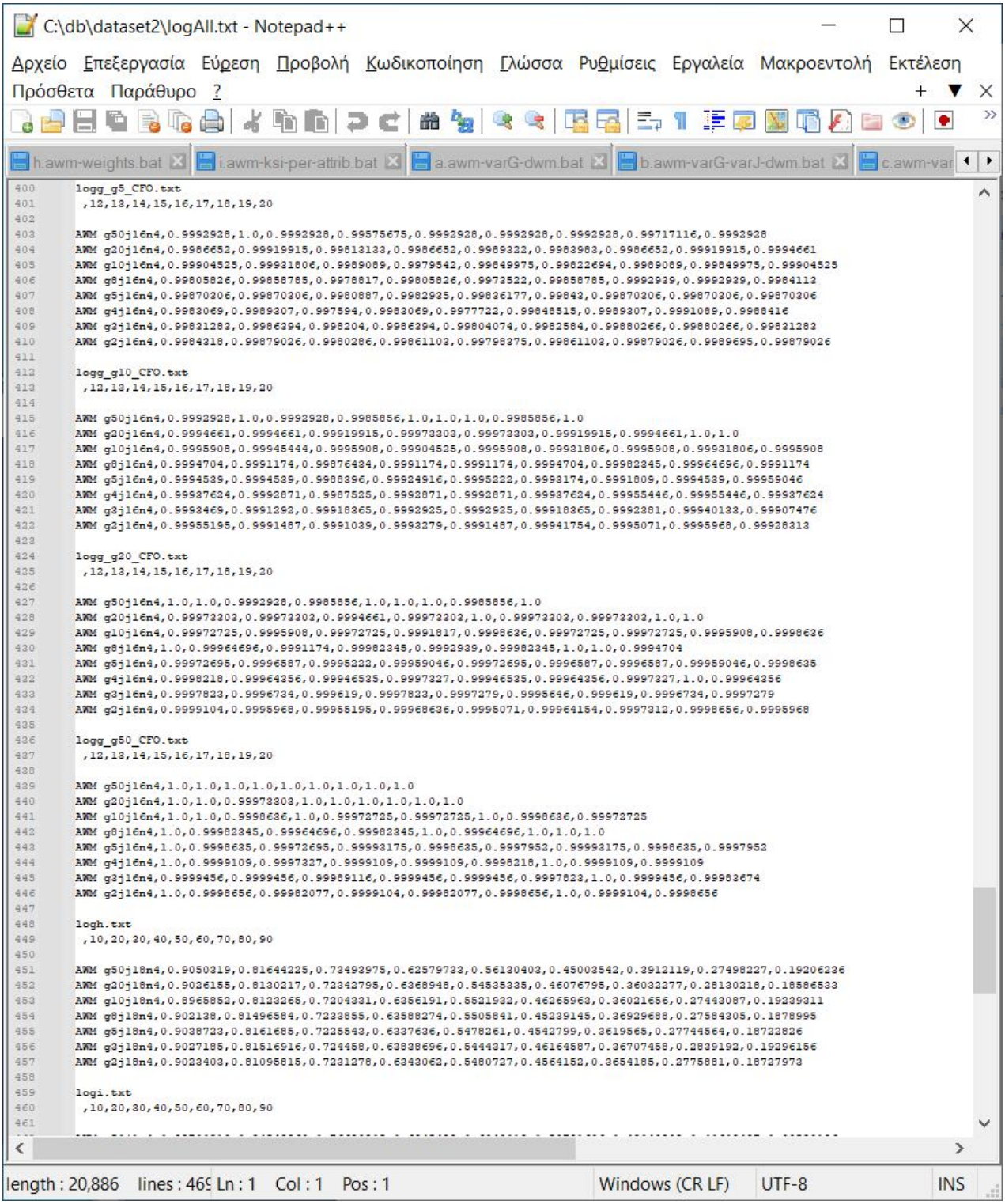
Εικόνα 41: Τα τρία datasets ακέραιων και δεκαδικών αριθμών



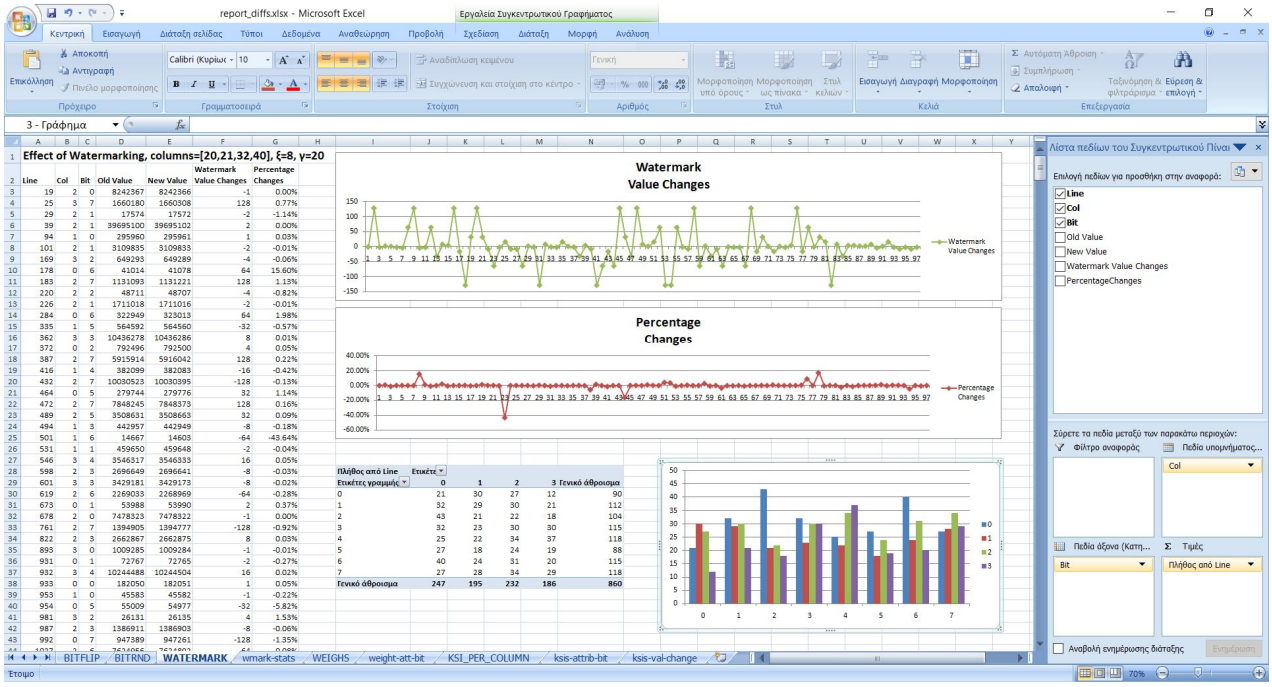
Εικόνα 42: Δημιουργία, συλλογή και οργάνωση γραφημάτων από τα σενάρια



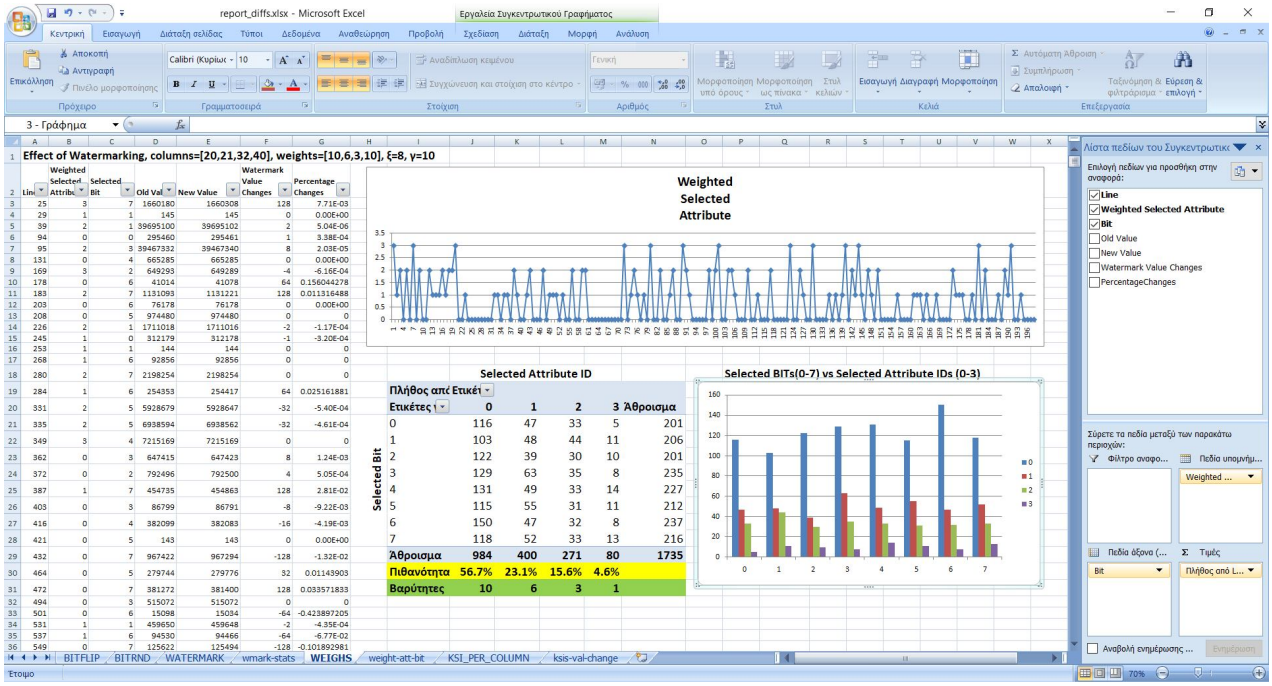
Εικόνα 43: Σενάριο δημιουργίας πολλαπλών υδατογραφήσεων και τύπων επιθέσεων



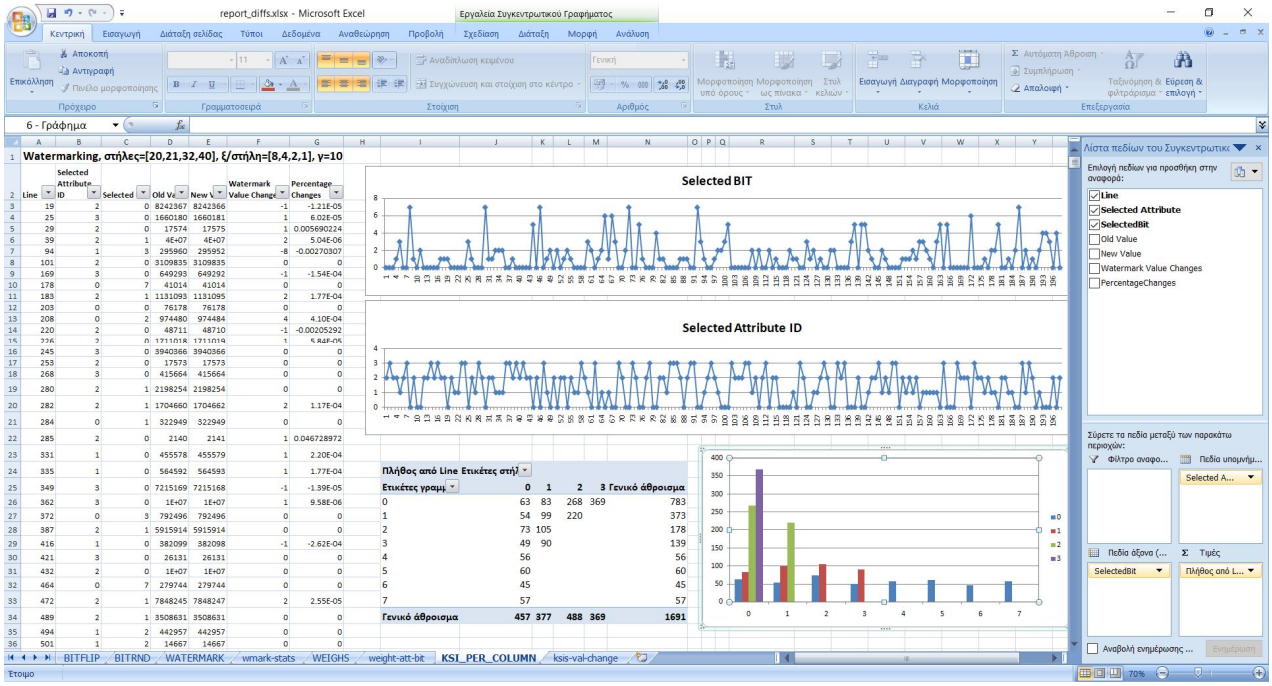
Εικόνα 44: Συγκεντρωτικό αρχείο Log από την εκτέλεση των σεναρίων



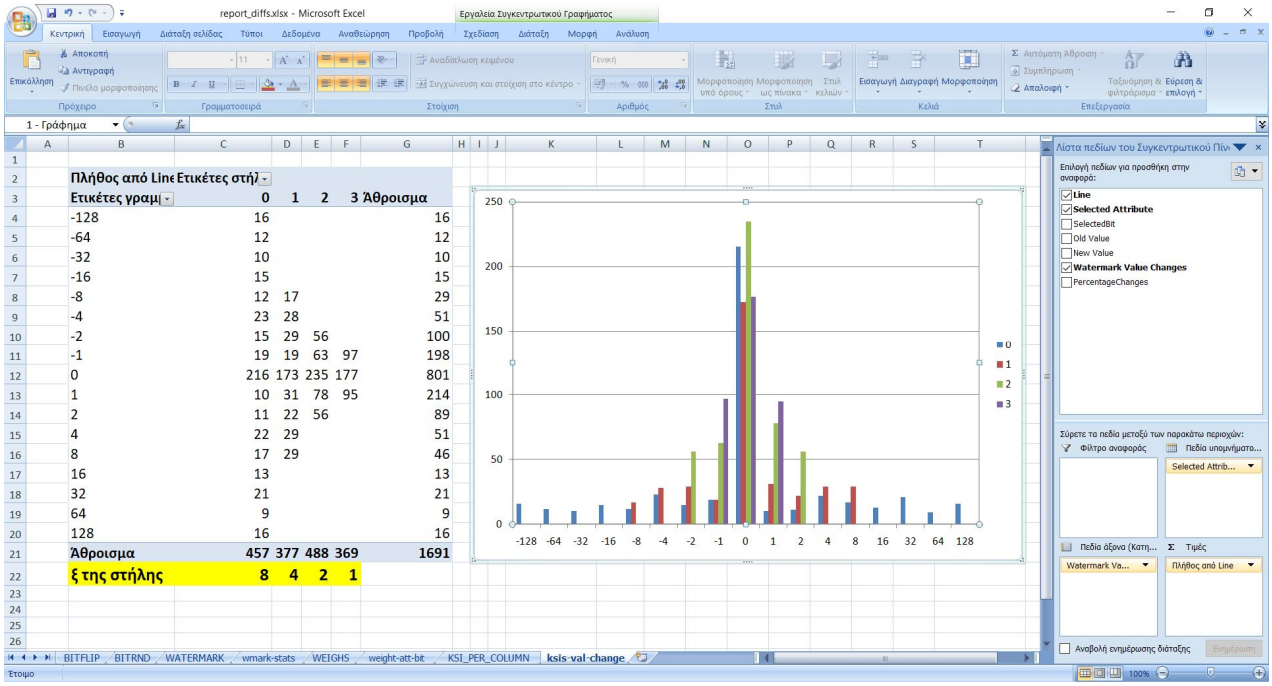
Εικόνα 45: Αλλοίωση των δεδομένων, των χαρακτηριστικών και των bits που επιφέρει μια υδατογράφηση



Εικόνα 46: Γραφήματα παραλλαγής του αλγόριθμου με διαφορετική πιθανότητα επιλογής χαρακτηριστικών



Εικόνα 47: Γραφήματα παραλλαγής του αλγόριθμου με διαφορετικά ξ σε κάθε χαρακτηριστικό



Εικόνα 48: Μεταβολή στις τιμές για διαφορετικά ξ σε κάθε χαρακτηριστικό