



ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ



**“Η ΕΠΙΧΕΙΡΗΜΑΤΙΚΗ ΝΟΗΜΟΣΥΝΗ ΚΑΙ ΟΙ ΕΦΑΡΜΟΓΕΣ
ΤΗΣ ΣΤΗΝ ΔΙΟΙΚΗΣΗ ΤΩΝ ΕΠΙΧΕΙΡΗΣΕΩΝ.”**

Κορμπάκη Όλγα
Αικατερίνη Βιδούρη
Σπυριδούλα Πολονύφη

Επιβλέπων Καθηγητής: Βλαχόπουλος Γιώργος

Πατρα 2009

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΕΡΙΕΧΟΜΕΝΑ.....Σελ 2

ΠΡΟΛΟΓΟΣ.....Σελ 10

ΜΕΡΟΣ Α

**Α1 Η ΕΝΝΟΙΑ ΤΩΝ ΠΛΗΡΟΦΟΡΙΑΚΩΝ
ΣΥΣΤΗΜΑΤΩΝ.....Σελ 11**

ΕΙΣΑΓΩΓΗ.....Σελ 11

Α1.1- Η ΕΝΝΟΙΑ ΤΟΥ ΣΥΣΤΗΜΑΤΟΣ.....Σελ 11

Α1.1.1- ΕΙΔΗ ΣΥΣΤΗΜΑΤΩΝ.....Σελ 12

Α1.1.2- ΠΑΡΑΔΕΙΓΜΑΤΑ ΣΥΣΤΗΜΑΤΩΝ.....Σελ 12

Α1.1.3- ΥΠΟΣΥΣΤΗΜΑΤΑ.....Σελ 13

Α1.1.4- ΕΙΣΡΟΗ, ΕΠΕΞΕΡΓΑΣΙΑ, ΕΚΡΟΗ,

ΠΕΡΙΒΑΛΛΟΝ.....Σελ 13-17

Α1.1.5- ΤΥΠΟΙ ΣΥΣΤΗΜΑΤΩΝ.....Σελ 17-18

**Α.1.2- ΔΕΔΟΜΕΝΑ, ΠΛΗΡΟΦΟΡΙΕΣ ΚΑΙ ΠΛΗΡΟΦΟΡΙΑΚΑ
ΣΥΣΤΗΜΑΤΑ ΔΙΟΙΚΗΣΗΣ.....Σελ 19.**

Α.1.2.1- ΓΕΝΙΚΑ ΠΕΡΙ ΠΛΗΡΟΦΟΡΙΑΚΩΝ

ΣΥΣΤΗΜΑΤΩΝ.....Σελ 19-21.

A.1.2.2- ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΠΛΗΡΟΦΟΡΙΑΣ.....Σελ 21-23

A.1.3.- ΤΟ ΓΝΩΣΤΙΚΟ ΠΕΔΙΟ ΤΩΝ ΠΛΗΡΟΦΟΡΙΑΚΩΝ

ΣΥΣΤΗΜΑΤΩΝ ΔΙΟΙΚΗΣΗΣ.....Σελ 23

A.1.3.1- ΟΙΚΟΝΟΜΙΚΕΣ ΕΠΙΣΤΗΜΕΣ.....Σελ 24

A.1.3.2 – ΕΠΙΧΕΙΡΗΣΙΑΚΗ ΈΡΕΥΝΑ.....Σελ 25

A.1.3.3- ΘΕΩΡΙΑ ΔΙΟΙΚΗΣΗΣ ΚΑΙ

ΟΡΓΑΝΩΣΗΣ.....Σελ 25

A.1.3.4- ΕΠΙΣΤΗΜΗ ΤΩΝ ΗΛΕΚΤΡΟΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ ΚΑΙ

ΠΛΗΡΟΦΟΡΙΚΗ.....Σελ 26

A2 ΓΕΝΙΚΗ ΘΕΩΡΗΣΗ ΚΑΙ ΚΑΤΑΤΑΞΗ ΠΛΗΡΟΦΟΡΙΑΚΩΝ

ΣΥΣΤΗΜΑΤΩΝ

ΔΙΟΙΚΗΣΗΣ..... Σελ 27

ΕΙΣΑΓΩΓΗ.....Σελ 27

A.2.1.-Η ΕΝΝΟΙΑ ΤΟΥ ΣΥΣΤΗΜΑΤΟΣ.....Σελ 27-29

A.2.2.- ΠΛΗΡΟΦΟΡΙΑΚΑ ΣΥΣΤΗΜΑΤΑ ΔΙΟΙΚΗΣΗΣ ΚΑΙ

ΟΡΓΑΝΩΤΙΚΗ ΔΟΜΗ ΕΠΙΧΕΙΡΗΣΗΣ.....Σελ 29

A.2.2.1- Η ΦΥΣΙΟΓΝΩΜΙΑ ΤΩΝ ΠΛΗΡΟΦΟΡΙΑΚΩΝ

ΣΥΣΤΗΜΑΤΩΝ ΔΙΟΙΚΗΣΗΣ.....Σελ 29-33

A.2.2.2- ΣΥΣΤΗΜΑ ΧΡΗΣΤΗ- ΥΠΟΛΟΓΙΣΤΗ.....Σελ 33

A.2.2.3- ΟΛΟΚΛΗΡΩΜΕΝΟ ΣΥΣΤΗΜΑ.....	Σελ 34
A.2.2.4- Η ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ.....	Σελ 36
A.2.2.5- ΧΡΗΣΙΜΟΠΟΙΗΣΗ ΜΟΝΤΕΛΩΝ.....	Σελ 37
A.2.3- Η ΕΞΕΛΙΞΗ ΤΩΝ ΠΛΗΡΟΦΟΡΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ ΔΙΟΙΚΗΣΗΣ.....	Σελ 38
A.2.3.1- ΠΣΔ ΚΑΙ ΣΥΣΤΗΜΑΤΑ ΔΙΑΧΕΙΡΙΣΗΣ ΠΟΡΩΝ ΠΛΗΡΟΦΟΡΙΩΝ.....	Σελ 39
A.2.3.2- ΠΛΗΡΟΦΟΡΙΑΚΟ ΣΥΣΤΗΜΑ ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΤΕΛΙΚΟΙ ΧΡΗΣΤΕΣ.....	Σελ 40
A.2.3.3- ΣΥΣΤΗΜΑΤΑ ΕΠΙΤΕΛΙΚΗΣ ΠΛΗΡΟΦΟΡΗΣΗΣ.....	Σελ 41
A.2.4- ΥΠΟΣΥΣΤΗΜΑΤΑ ΕΝΟΣ ΠΛΗΡΟΦΟΡΙΑΚΟΥ ΣΥΣΤΗΜΑΤΟΣ ΔΙΟΙΚΗΣΗΣ.....	Σελ 42
A.2.4.1- ΥΠΟΣΥΣΤΗΜΑΤΑ ΛΕΙΤΟΥΡΓΙΩΝ.....	Σελ 43
A.2.4.2- ΥΠΟΣΥΣΤΗΜΑΤΑ ΟΡΓΑΝΩΤΙΚΩΝ ΕΠΙΠΕΔΩΝ.....	Σελ 44
A.2.5- ΦΥΣΙΚΗ ΔΟΜΗ ΕΝΟΣ ΠΛΗΡΟΦΟΡΙΑΚΟΥ ΣΥΣΤΗΜΑΤΟΣ ΔΙΟΙΚΗΣΗΣ.....	Σελ 45
A3 ΟΡΓΑΝΩΣΗ ΔΕΔΟΜΕΝΩΝ.....	Σελ 47

ΕΙΣΑΓΩΓΗ.....	Σελ 47
Α.3.1- ΣΥΜΒΑΤΙΚΕΣ ΜΕΘΟΔΟΙ ΟΡΓΑΝΩΣΗΣ	
ΔΕΔΟΜΕΝΩΝ.....	Σελ 47
Α.3.1.1- ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ.....	Σελ 47-51
Α.3.1.2- ΤΥΠΟΙ ΛΟΓΙΚΩΝ ΕΓΓΡΑΦΩΝ.....	Σελ 51
Α.3.1.3- ΣΕΙΡΙΑΚΗ ΟΡΓΑΝΩΣΗ ΑΡΧΕΙΩΝ.....	Σελ 52
Α.3.1.4- ΟΡΓΑΝΩΣΗ ΑΡΧΕΙΩΝ ΆΜΕΣΗΣ	
ΠΡΟΣΠΕΛΑΣΗΣ.....	Σελ 54
Α.3.1.4.1- ΜΕΘΟΔΟΙ ΣΥΣΧΕΤΙΣΜΟΥ ΚΛΕΙΔΙΟΥ ΚΑΙ	
ΔΙΕΥΘΥΝΣΗΣ.....	Σελ 55-57
Α.3.1.4.2- ΕΝΗΜΕΡΩΣΗ ΑΡΧΕΙΟΥ ΆΜΕΣΗΣ	
ΠΡΟΣΠΕΛΑΣΗΣ.....	Σελ 57-58
Α.3.1.4.3- ΑΝΑΛΗΨΗ ΠΛΗΡΟΦΟΡΙΩΝ ΑΠΟ ΤΟ	
ΑΡΧΕΙΟ.....	Σελ 59
Α.3.2.- ΣΥΝΘΕΤΟΙ ΜΕΘΟΔΟΙ ΟΡΓΑΝΩΣΗΣ	
ΑΡΧΕΙΩΝ.....	Σελ 59
Α.3.2.1.- ΟΡΓΑΝΩΣΗ ΚΑΤΑΛΟΓΟΥ(LIST	
ORGANIZATION).....	Σελ 60-63
Α.3.2.2.- ΠΙΝΑΚΟΠΟΙΗΜΕΝΗ ΣΕΙΡΙΑΚΗ	
ΟΡΓΑΝΩΣΗ(INDEXED SEQUENTIAL ACCESS METHOD-	
ISAM).....	Σελ 63

A.3.3.- ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ.....	Σελ 64
A.3.3.1.- ΣΥΣΤΗΜΑΤΑ ΔΙΑΧΕΙΡΙΣΗΣ ΒΑΣΕΩΝ ΔΕΔΟΜΕΝΩΝ.....	Σελ 65-69
A.3.3.2.- ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΣΥΣΤΗΜΑΤΩΝ ΒΑΣΕΩΝ ΔΕΔΟΜΕΝΩΝ.....	Σελ 69-72.
A.3.3.3.- ΔΟΜΕΣ ΒΑΣΕΩΝ ΔΕΔΟΜΕΝΩΝ.....	Σελ 72-75

ΜΕΡΟΣ Β

B4 ΛΗΨΗ ΑΠΟΦΑΣΕΩΝ.....	Σελ 76
ΕΙΣΑΓΩΓΗ.....	Σελ 76
B4.1- ΓΕΝΙΚΗ ΜΕΘΟΔΟΛΟΓΙΑ ΛΗΨΗΣ ΑΠΟΦΑΣΗΣ.....	Σελ 77
B4.1.1- ΚΑΘΟΡΙΣΜΟΣ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ.....	Σελ 78
B4.1.2- ΕΝΑΛΛΑΚΤΙΚΕΣ ΛΥΣΕΙΣ.....	Σελ 79-81
B4.1.3- ΑΞΙΟΛΟΓΗΣΗ ΛΥΣΕΩΝ.....	Σελ 82
B4.1.4- ΕΠΙΛΟΓΗ ΤΗΣ ΛΥΣΗΣ.....	Σελ 83
B4.2- ΔΙΑΔΙΚΑΣΙΑ ΛΗΨΗΣ ΑΠΟΦΑΣΗΣ.....	Σελ 84-91
B4.3- " ΠΑΙΓΝΙΑ ".....	Σελ 92-94
B4.4- ΤΥΠΟΙ ΚΑΙ ΚΑΤΗΓΟΡΙΕΣ ΑΠΟΦΑΣΕΩΝ.....	Σελ 94

B4.4.1- ΤΥΠΟΙ ΑΠΟΦΑΣΕΩΝ.....	Σελ 94
B4.4.2- ΥΠΟΚΕΙΜΕΝΙΚΟΤΗΤΑ ΚΑΙ ΑΠΟΦΑΣΕΙΣ.....	Σελ 97-98
B4.5- ΜΟΝΤΕΛΑ ΑΠΟΦΑΣΕΩΝ.....	Σελ 99-102.
B4.6- ΛΗΨΗ ΑΠΟΦΑΣΕΩΝ ΚΑΙ ΠΛΗΡΟΦΟΡΙΑΚΑ ΣΥΣΤΗΜΑΤΑ.....	Σελ 102
B4.6.1- ΔΙΑΔΙΚΑΣΙΑ ΛΗΨΗΣ ΑΠΟΦΑΣΗΣ ΚΑΙ ΠΛΗΡΟΦΟΡΙΑΚΑ ΣΥΣΤΗΜΑΤΑ.....	Σελ 102
B4.6.2- ΟΡΓΑΝΩΣΗ ΚΑΙ ΠΛΗΡΟΦΟΡΙΑΚΑ ΣΥΣΤΗΜΑΤΑ.....	Σελ 103-105
B4.6.3- ΑΠΟΦΑΣΕΙΣ ΚΑΙ ΣΧΕΔΙΑΣΜΟΣ ΠΛΗΡΟΦΟΡΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ.....	Σελ 105-107
 B5 ΟΙ ΤΕΧΝΙΚΕΣ ΕΞΟΡΥΞΗΣ ΑΠΟ ΔΕΔΟΜΕΝΑ.....	
B5.1- ΕΙΣΑΓΩΓΗ.....	Σελ 108-109
B5.2- ΚΑΝΟΝΕΣ ΣΥΣΧΕΤΙΣΗΣ.....	Σελ 110
B5.2.1- ΓΕΝΙΚΑ.....	Σελ 110
B5.2.2- ΑΛΓΟΡΙΘΜΟΙ ΣΥΣΧΕΤΙΣΗΣ.....	Σελ 112
B5.2.3. - ΔΙΑΔΟΧΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ(SEQUENTIAL ALGORITHMS).....	Σελ 112
B5.2.3.1 - ΠΑΡΑΛΛΗΛΟΙ ΑΛΓΟΡΙΘΜΟΙ(PARALLEL ALGORITHMS).....	Σελ 115

B5.2.3.2 - ΠΟΣΟΤΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ(QUANTITATIVE ALGORITHMS)	Σελ 116
B5.2.3.3 - ΓΕΝΙΚΕΥΜΕΝΟΙ ΑΛΓΟΡΙΘΜΟΙ(GENERALIZED ALGORITHMS)	Σελ 117
B5.2.3.4 - ΧΩΡΙΚΟΙ ΚΑΙ ΧΡΟΝΟΛΟΓΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ(SPATIAL AND TEMPORAL)	Σελ 117
B5.2.5 - ΔΙΑΦΟΡΑ ΘΕΜΑΤΑ	Σελ 118
B5.3- ΚΑΝΟΝΕΣ ΤΑΞΙΝΟΜΗΣΗΣ	Σελ 119
B5.3.1- ΓΕΝΙΚΑ	Σελ 119-120
B5.3.2- ΑΛΓΟΡΙΘΜΟΙ ΤΑΞΙΝΟΜΗΣΗΣ ΔΕΝΔΡΩΝ	Σελ120-122
B5.3.3- ΚΑΝΟΝΕΣ ΔΙΑΚΛΑΔΩΣΗΣ / ΔΙΑΙΡΕΣΗΣ ΔΕΝΔΡΩΝ(SPLITTING RULES)	Σελ 122-125
B5.3.4- ΜΕΘΟΔΟΙ ΠΕΡΙΟΡΙΣΜΟΥ ΈΚΤΑΣΗΣ ΤΩΝ ΔΕΝΤΡΩΝ	Σελ 125-127
B5.3.5- ΣΥΓΚΡΙΣΗ ΜΕ ΑΛΛΕΣ ΜΕΘΟΔΟΛΟΓΙΕΣ	Σελ 127
B5.4.- ΚΑΝΟΝΕΣ ΟΜΑΔΟΠΟΙΗΣΗΣ	Σελ 129
B5.4.1- ΓΕΝΙΚΑ	Σελ 129-133
B5.4.2- ΑΛΓΟΡΙΘΜΟΙ ΟΜΑΔΟΠΟΙΗΣΗΣ	Σελ 133
B5.4.2.1- ΔΙΑΧΩΡΙΣΤΙΚΗ ΟΜΑΔΟΠΟΙΗΣΗ	Σελ 133-135
B5.4.2.2- ΙΕΡΑΡΧΙΚΗ ΟΜΑΔΟΠΟΙΗΣΗ	Σελ 135

B5.4.2.3- ΟΜΑΔΟΠΟΙΗΣΗ ΒΑΣΗ ΣΥΝΘΗΚΩΝ	
ΠΥΚΝΟΤΗΤΑΣ.....	Σελ 138
B5.4.2.4- ΟΜΑΔΟΠΟΙΗΣΗ ΒΑΣΗ ΠΛΕΓΜΑΤΩΝ.....	Σελ 139
B5.4.2.5- ΟΜΑΔΟΠΟΙΗΣΗ FUZZY.....	Σελ 139
B5.4.2.6- ΟΜΑΔΟΠΟΙΗΣΗ ΒΑΣΗ ΚΑΤΑΝΟΜΩΝ	
GAUSSIAN.....	Σελ 140
B5.4.3- ΔΙΑΦΟΡΑ ΘΕΜΑΤΑ.....	Σελ 140-141
ΣΥΜΠΕΡΑΣΜΑΤΑ.....	Σελ 142
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	Σελ 143-145

ΠΡΟΛΟΓΟΣ

Η Πτυχιακή αυτή, πραγματεύεται την Επιχειρησιακή Νοημοσύνη (**Business Intelligence**) και τις εφαρμογές αυτής στην Διοίκηση των Επιχειρήσεων.

Η Επιχειρησιακή Νοημοσύνη σχετίζεται με την αρωγή και ενδοκίμηση του συνδυασμού τεχνολογίας . εφαρμογών- προγραμμάτων και τις πρακτικές για την συλλογή . τον συσχετισμό και αλληλεπίδραση καθώς και την ανάλυση της πληροφορίας αυτής καθ' εαυτής. Σκοπός της Επιχειρησιακής Νοημοσύνης είναι η υποστήριξη καλύτερων επιχειρηματικών αποφάσεων.

Αναλυτικότερα, θα γίνει εκτενής αναφορά στην Έννοια των Πληροφοριακών Συστημάτων, στην Γενική θεώρηση και κατάταξη αυτών, καθώς επίσης και στην Οργάνωση δεδομένων για Αποτελεσματική Αποθήκευση και Επεξεργασία. Στη συνέχεια , θα αναφερθούμε στο ρόλο της Λήψης Αποφάσεων η οποία είναι ένα από τα βασικά εργαλεία για τη διοίκηση της επιχείρησης.

Απαραίτητη βέβαια, κρίνεται μια αναλυτική αναφορά σχετικά με τις τεχνικές Εξόρυξης από Δεδομένα τα οποία αποτελούν ένα από τα σημαντικότερα εργαλεία ανάλυσης δεδομένων των τελευταίων δεκαετιών. Τέλος, η ανάλυση του Πληροφοριακού Συστήματος της Επιχείρησης και Περιβάλλον καθώς και τα Στάδια Εξέλιξης του Πληροφοριακού Συστήματος της Επιχείρησης συνθέτουν την εικόνα του θέματος που αναλύουμε γενικότερα

ΚΕΦΑΛΑΙΟ 1ο

Η ΕΝΝΟΙΑ ΤΩΝ ΠΛΗΡΟΦΟΡΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΕΙΣΑΓΩΓΗ

Στο κεφάλαιο αυτό παρουσιάζεται η έννοια του “Συστήματος” και αναπτύσσονται τα βασικά συστατικά και οι κατηγορίες των Συστημάτων. Παρουσιάζονται επίσης οι έννοιες των δεδομένων, των πληροφοριών και της επεξεργασίας. Τέλος δίνεται η έννοια του “Πληροφοριακού Συστήματος Διοίκησης”. Ο κύριος σκοπός του κεφαλαίου είναι η κατανόηση των παραπάνω εννοιών.

1.1. Η ΕΝΝΟΙΑ ΤΟΥ ΣΥΣΤΗΜΑΤΟΣ

Με τον όρο Σύστημα νοείται ένα σύνολο στοιχείων ,διαρθρωμένων με κάποια συγκεκριμένη οργανωτική δομή που επιτελεί μια σειρά δραστηριοτήτων και επιδιώκει την επίτευξη ενός προκαθορισμένου σκοπού.

Ο ορισμός αυτός υπονοεί ότι όλοι οι παράγοντες που σχετίζονται με ένα σύστημα ,τόσο σε επίπεδο στοιχείων όσο και σε επίπεδο συνόλου, είναι καθορισμένοι και κανείς δεν αφήνεται στην τύχη. Έτσι, τα στοιχεία που απαρτίζουν το σύστημα είναι σαφώς καθορισμένα, όπως επίσης καθορισμένη είναι η λειτουργία του καθενός από αυτά και ο επιμέρους αντικειμενικός σκοπός του. Επίσης η αλληλεξάρτηση, η αλληλεπίδραση και η συνοχή όλων των στοιχείων του συστήματος είναι καθορισμένη από ένα σχέδιο εσωτερικής οργάνωσης και δομής με κάποιο συγκεκριμένο τελικό αντικειμενικό σκοπό σε επίπεδο συνόλου.

1.1.1 ΕΙΔΗ ΣΥΣΤΗΜΑΤΩΝ

Τα συστήματα μπορούν να χωριστούν σε **(3)** κατηγορίες:

ΦΥΣΙΚΑ ΣΥΣΤΗΜΑΤΑ:

Ένας οποιοσδήποτε ζωντανός (φυσικός) οργανισμός αποτελεί φυσικό σύστημα.

ΤΕΧΝΗΤΑ ΣΥΣΤΗΜΑΤΑ:

Ένα μηχάνημα που αποτελείται μόνο από υλικά μέσα θεωρείται τεχνητό σύστημα.

ΜΕΙΚΤΑ ΣΥΣΤΗΜΑΤΑ:

Πρόκειται τα συστήματα των οποίων τα στοιχεία είναι άνθρωποι και υλικά μέσα.

1.1.2 ΠΑΡΑΔΕΙΓΜΑΤΑ ΣΥΣΤΗΜΑΤΩΝ

Ένας Οργανισμός Παραγωγής Προϊόντων (Βιομηχανική Μονάδα) αποτελεί Σύστημα του οποίου τα στοιχεία είναι άνθρωποι και υλικά μέσα (μεικτό σύστημα). Το ίδιο ισχύει για ένα Οργανισμό Παροχής Υπηρεσιών (Νοσοκομείο, Δημοτικό Κατάστημα,...). Επίσης ένα Πληροφοριακό Σύστημα είναι ένα μεικτό σύστημα διότι σ' αυτό ένα σημαντικό στοιχείο του είναι οι χρήστες.

Ο άνθρωπος σαν οντότητα αποτελεί την τελειότερη μορφή φυσικού συστήματος.

Ένα ρομπότ ή ένας Ηλεκτρονικός Υπολογιστής είναι ένα τεχνητό σύστημα.

1.1.3 ΥΠΟΣΥΣΤΗΜΑΤΑ

Μια προσεκτική μελέτη των παραπάνω παραδειγμάτων οδηγεί στο συμπέρασμα ότι καθένα από αυτά αποτελείται από επιμέρους τμήματα (Υποσυστήματα) με την απαραίτητη αυτοτέλεια και χαρακτηριστικά συστήματος, πού υπόκεινται όμως και εξυπηρετούν τους ευρύτερους στόχους του Συστήματος στο οποίο ανήκουν.

Το Λογιστήριο του Νοσοκομείου, η Τεχνική Υπηρεσία του Δήμου, το Πεπτικό Σύστημα του Ανθρώπου και η Κεντρική Μονάδα Επεξεργασίας του Υπολογιστή αποτελούν υποσυστήματα των Συστημάτων που αναφέρθηκαν παραπάνω σαν παραδείγματα.

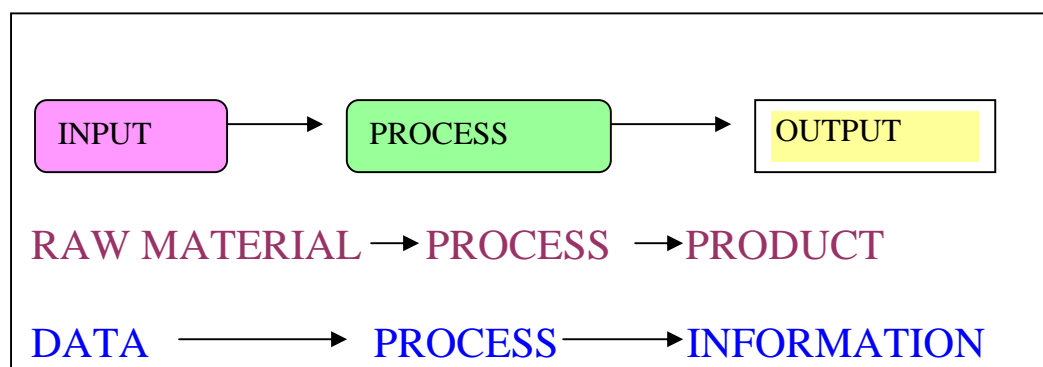
Κάθε υποσύστημα τέλος μπορεί να αποτελεί σύστημα για τα οργανικά μέρη που περιλαμβάνει. Για παράδειγμα, το τμήμα Οδοποιίας αποτελεί Υποσύστημα της Τεχνικής Υπηρεσίας του Δήμου, η Μονάδα Ελέγχου Υποσύστημα της Κεντρικής Μονάδας Επεξεργασίας, το στομάχι Υποσύστημα του Πεπτικού Συστήματος.

1.1.4 ΕΙΣΡΟΗ, ΕΠΕΞΕΡΓΑΣΙΑ, ΕΚΡΟΗ, ΠΕΡΙΒΑΛΛΟΝ

Κάθε σύστημα επικοινωνεί με το περιβάλλον του δεχόμενο εισροές από αυτό, τις οποίες μετασχηματίζει στο εσωτερικό του και αποδίδει με την σειρά του τα αποτελέσματα του μετασχηματισμού αυτού στο περιβάλλον, παράγοντας έτσι εκροές προς αυτό. Το σύνολο των εισροών προς το σύστημα(αλλά και η λειτουργία της εισροής)θα αναφέρεται στο εξής με τον όρο **είσοδος (INPUT)**. Αντίστοιχα, το σύνολο των εκροών από το σύστημα(αλλά και η λειτουργία της εκροής)θα αναφέρεται με τον όρο **έξοδος (OUTPUT)**. Τέλος, η διαδικασία

του μετασχηματισμού των εισροών στο εσωτερικό του συστήματος θα καλείται **επεξεργασία(PROCESS)**.

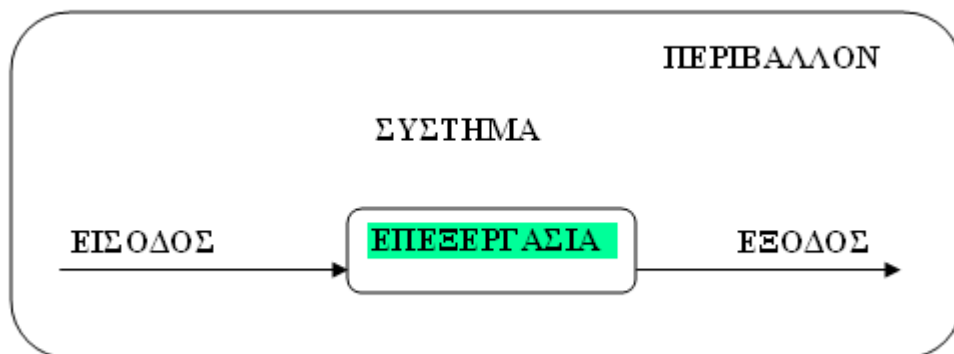
Για μια Μονάδα Παραγωγής η είσοδος μεταξύ των άλλων περιλαμβάνει τις πρώτες ύλες και η έξοδος το τελικό προϊόν, ενώ για τον Ηλεκτρονικό Υπολογιστή η είσοδος είναι τα δεδομένα και η έξοδος οι πληροφορίες που παράγει για το χρήστη που τον χρησιμοποιεί.



Σχήμα 1-1:Είσοδος-Επεξεργασία-Έξοδος

Εκτός από την τυποποιημένη διαδικασία εισόδου, επεξεργασίας, εξόδου, ένα σύστημα, στο βαθμό που αποτελεί μέρος ενός ευρύτερου Συστήματος, δέχεται άτυπες, μη προκαθορισμένες εισροές από το περιβάλλον του, που αποτελούν παράγοντα επιβίωσης και προσαρμογής σε νέες θέσεις ισορροπίας. Οι Μονάδες Παραγωγής για παράδειγμα δεν είναι δυνατόν να επιβιώσουν για μεγάλο χρονικό διάστημα αν δεν έχουν ευέλικτους μηχανισμούς προσαρμογής στις διαμορφούμενες συνθήκες(Νέες αγορές, Νέα Τεχνολογία, Νομοθετικό Πλαίσιο,...)Το ίδιο ισχύει και για τον ανθρώπινο οργανισμό του οποίου η ικανότητα επιβίωσης σε μεγάλο βαθμό εξαρτάται από την ικανότητα προσαρμογής του στο περιβάλλον.

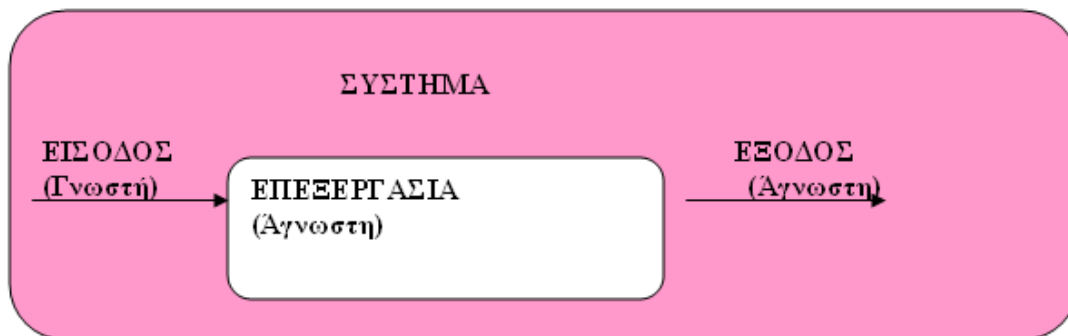
Στο Σχήμα 1-2 φαίνεται απλοποιημένα η εικόνα ενός συστήματος και η σύνδεσή του με το περιβάλλον του.



Σχήμα 1-2: Απλοποιημένη Εικόνα Συστήματος

Ένα σύστημα είναι πλήρως γνωστό στον χρήστη του όταν έχει αποσαφηνιστεί η είσοδος προς αυτό, η επιτελούμενη επεξεργασία στο εσωτερικό του, καθώς και η παραγόμενη από αυτό έξοδος. Ωστόσο, συμβαίνει συχνά να χρησιμοποιεί κάποιος ένα σύστημα γνωρίζοντας μόνο τι εισέρχεται σ' αυτό και τι εξέρχεται από αυτό χωρίς όμως να γνωρίζει το παραμικρό ή να γνωρίζει πολύ λίγα σχετικά με το τι συμβαίνει στο εσωτερικό του. Ένα τέτοιο σύστημα χαρακτηρίζεται με τον όρο **μαύρο κουτί (BLACK BOX)**. Χαρακτηριστικό παράδειγμα μαύρου κουτιού είναι η περίπτωση του χρήστη ενός έτοιμου προγράμματος μισθοδοσίας προσωπικού, ο οποίος γνωρίζει την είσοδο των μισθοδοτικών στοιχείων προς αυτό (από το εγχειρίδιο που συνοδεύει το πρόγραμμα ή καθοδηγούμενος από το ίδιο το πρόγραμμα κατά την εκτέλεσή του,) γνωρίζει τα αποτελέσματα που μπορεί να πάρει από το πρόγραμμα (καταστάσεις κλπ.), αλλά δεν γνωρίζει τον τρόπο με τον οποίο το πρόγραμμα επεξεργάζεται τα εισαχθέντα μισθοδοτικά στοιχεία.

Στο παρακάτω Σχήμα 1-3 φαίνεται σχηματικά η εικόνα ενός συστήματος μαύρου κουτιού.



Σχήμα 1-3: Σύστημα μαύρο κουτί

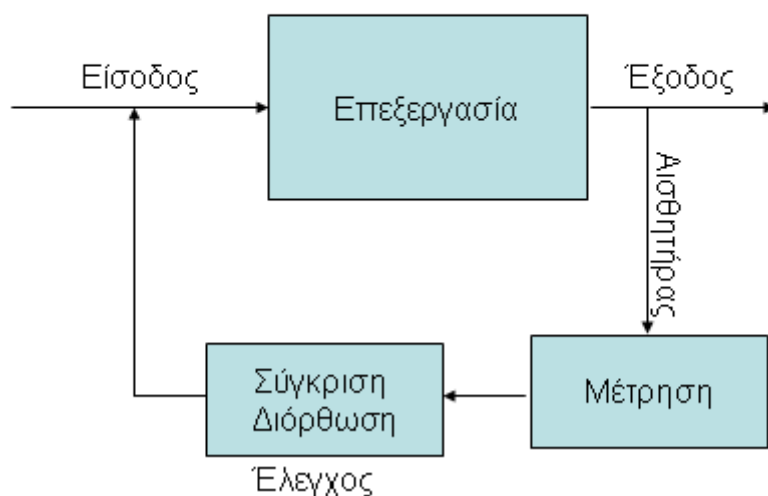
Προκειμένου να εξασφαλιστεί ότι ένα σύστημα που σχεδιάστηκε και λειτουργεί για την επίτευξη ενός αντικειμενικού σκοπού δεν απομακρύνεται από το σκοπό αυτό, ελέγχεται συνεχώς ως προς την παραγόμενη έξοδο του. Κατά τον έλεγχο αυτό η έξοδος που παράγει το σύστημα μετριέται και συγκρίνεται προς τα πρότυπα που έχουν τεθεί σύμφωνα με τον σχεδιασμό του συστήματος. Εάν η σύγκριση δείξει απόκλιση από τα πρότυπα τότε αναλαμβάνονται διορθωτικές ενέργειες. Για παράδειγμα, ας υποθέσουμε ότι μια εταιρεία τυποποίησης τροφίμων έχει εγκαταστήσει ένα σύστημα συσκευασίας ζάχαρης σε πακέτα του ενός κιλού και ότι οι ισχύουσες αγορανομικές διατάξεις δεν επιτρέπουν αποκλίσεις μεγαλύτερες των 5 γραμμαρίων. Γίνεται φανερό ότι το σύστημα αυτό πρέπει να υφίσταται οπωσδήποτε συνεχή έλεγχο ως προς το βάρος των πακέτων που παράγει.

Εάν τα βάρη αυτά αποκλίνουν πέρα από τα 5 γραμμάρια ,τότε πρέπει να αναληφθεί διορθωτική ενέργεια στο σύστημα με επέμβαση, για παράδειγμα, στη μηχανή συσκευασίας.

Συχνά ο έλεγχος του συστήματος αυτοματοποιείται με εφοδιασμό του συστήματος με κατάλληλα υποσυστήματα που αναλαμβάνουν το έργο αυτό. Συνήθως απαιτείται ένα **αισθητήριο** υποσύστημα (**SENSOR**) που ευαισθητοποιείται από την παραγόμενη έξοδο και την μετρά. Οι μετρήσεις

αυτές διαβιβάζονται σε ένα άλλο υποσύστημα **ελέγχου (CONTROL)**, το οποίο εκτελεί σύγκριση προς τα πρότυπα και διαβιβάζει κατάλληλα διορθωτικά μηνύματα στο σύστημα. Η μέθοδος αυτή καλείται **έλεγχος με ανατροφοδότηση (FEEDBACK CONTROL)** και φαίνεται στο Σχήμα 1-4.

Σύστημα με Ανατροφοδότηση



Σχήμα 1-4: Σύστημα με ανατροφοδότηση

1.1.5 ΤΥΠΟΙ ΣΥΣΤΗΜΑΤΩΝ

Με βάση τα παραπάνω χαρακτηριστικά, τα Συστήματα κατατάσσονται σε πέντε κατηγορίες, ενώ δεν αποκλείεται κάποιο σύστημα να ανήκει και σε περισσότερες από αυτές.

1.ΑΙΤΙΟΚΡΑΤΙΚΑ ΣΥΣΤΗΜΑΤΑ: Χαρακτηρίζονται αυτά που οι εκροές μας δίνουν βέβαιη πρόβλεψη γεγονότων και τα στοιχεία τους είναι σε αυστηρά καθορισμένη σχέση μεταξύ τους. Η μηχανή του αυτοκινήτου, ο Ηλεκτρονικός Υπολογιστής, τα βιομηχανικά ρομπότ ανήκουν στη κατηγορία των Αιτιοκρατικών Συστημάτων.

Σε ένα Αιτιοκρατικό Σύστημα η κατάσταση του **S** σε μια δεδομένη χρονική στιγμή **t** είναι συνάρτηση του χρόνου **t** και της εισόδου **I** σε αυτό: **S=f(t,I)**

2. ΠΙΘΑΝΟΣΥΣΤΗΜΑΤΑ: Σε αντίθεση με τα Αιτιοκρατικά, τα Πιθανοσυστήματα παρέχουν αβέβαιη πρόβλεψη γεγονότων. Τα φυσικά και μεικτά συστήματα ανήκουν σε αυτή τη κατηγορία. Η ακριβής πρόβλεψη γεγονότων σε ένα ανθρώπινο οργανισμό, ασθένεια κλπ., είναι αδύνατη, όπως αδύνατη είναι και η πρόβλεψη γεγονότων σε ένα Εμπορικό Κατάστημα τη στιγμή που μέσα σε αυτό βρίσκονται άνθρωποι με άγνωστες διαθέσεις και προθέσεις.

3. ΚΛΕΙΣΤΑ ΣΥΣΤΗΜΑΤΑ: Είναι τα Συστήματα τα οποία έρχονται σε επαφή με το περιβάλλον τους μόνο μέσα από την τυπική διαδικασία εισόδου-εξόδου που τους εξασφαλίζει όμως μόνο μια σύντομη διάρκεια ζωής.

Το μέχρι το **1990** μεταπολεμικό καθεστώς της Αλβανίας αποτελεί τυπική περίπτωση κλειστού κοινωνικοπολιτικού Συστήματος το οποίο επί δεκαετίες λειτούργησε στη βάση συγκεκριμένων διπλωματικών και εμπορικών σχέσεων με άλλες χώρες αρνούμενο την οποιαδήποτε πολιτιστική επιρροή και επίδραση.

4. ΑΝΟΙΚΤΑ ΣΥΣΤΗΜΑΤΑ: Σε αντίθεση με τα κλειστά, τα Ανοικτά Συστήματα έχουν άτυπη αλληλεπίδραση με το περιβάλλον τους, προσαρμοζόμενα σε νέα δεδομένα και απαιτήσεις εξασφαλίζοντας μεγαλύτερο κύκλο ζωής. Τα Πληροφοριακά Συστήματα Διοίκησης που αποτελούν αντικείμενο αυτού του δοκιμίου, στο βαθμό μάλιστα που χρησιμοποιούν Τεχνολογία Ηλεκτρονικών Υπολογιστών με τη γνωστή εκρηκτική εξέλιξη, ανήκουν στη κατηγορία των Ανοικτών Συστημάτων .

5. ΕΥΣΤΑΘΗ ΣΥΣΤΗΜΑΤΑ: Χαρακτηρίζονται αυτά στα οποία γίνεται τακτικός έλεγχος της παραγόμενης εκροής. Τα Συστήματα Παραγωγής ανήκουν σε αυτήν την κατηγορία Συστημάτων.

1.2. ΔΕΔΟΜΕΝΑ, ΠΛΗΡΟΦΟΡΙΕΣ ΚΑΙ ΠΛΗΡΟΦΟΡΙΑΚΑ ΣΥΣΤΗΜΑΤΑ ΔΙΟΙΚΗΣΗΣ

1.2.1 ΓΕΝΙΚΑ ΠΕΡΙ ΠΛΗΡΟΦΟΡΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

Ας θεωρήσουμε ένα Δημοσιογραφικό Οργανισμό, ο οποίος με διάφορους μηχανισμούς, ανθρώπους και μέσα έχει σαν στόχο την πληροφόρηση του κοινού (**OUTPUT**), καταγράφοντας, αξιολογώντας και ερμηνεύοντας γεγονότα (**INPUT**). Ένα τέτοιο σύστημα θα μπορούσε να χαρακτηριστεί σαν Πληροφοριακό σύστημα, όπως θα μπορούσε να χαρακτηριστεί σαν οικονομικό σύστημα μια οποιαδήποτε παραγωγική μονάδα μετατροπής πρώτης ύλης σε τελικό προϊόν.

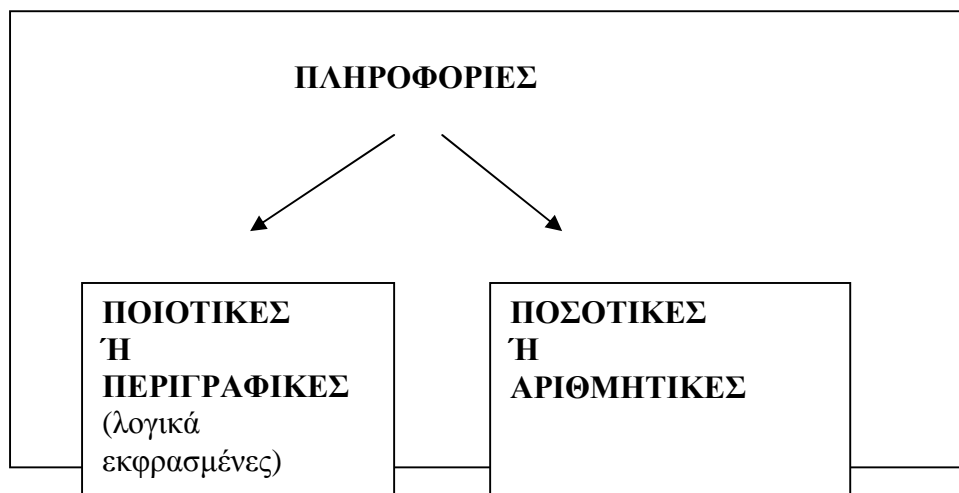
Ας θεωρήσουμε τώρα ένα Πληροφοριακό Σύστημα το οποίο, σαν υποσύστημα ενός οργανισμού, έχει σαν στόχο να παράγει πληροφορίες στα όργανα διοίκησης, επεξεργαζόμενο διάφορα δεδομένα, με σκοπό να υποστηρίζει διοικητικές πράξεις και αποφάσεις για την αποτελεσματικότερη άσκηση των καθηκόντων τους.

Ένα τέτοιο σύστημα ονομάζεται Πληροφοριακό Σύστημα Διοίκησης (Π.Σ.Δ., αγγλικός όρος **MANAGEMENT INFORMATION SYSTEM** ή εν συντομία **M.I.S**). Οι εισροές σε ένα τέτοιο σύστημα είναι δεδομένα (**DATA**) και οι εκροές πληροφορίες (**INFORMATION**).

Από αυτό φαίνεται ότι υπάρχει μια σημαντική διαφορά μεταξύ δεδομένων και πληροφοριών. Η σχέση τους είναι ανάλογη με αυτήν που υπάρχει μεταξύ πρώτων υλών και τελικού προϊόντος. Όπως το αλεύρι, η μαγιά, το νερό, το αλάτι κ.λπ. με κατάλληλη χημική και μηχανική κατεργασία μετατρέπονται σε ψωμί, έτσι και η εξέλιξη των τιμών των καταναλωτικών προϊόντων μας δίνει τον ρυθμό εξέλιξης του πληθωρισμού μετά από μια λογική και μηχανική ή ηλεκτρονική επεξεργασία. Σ' αυτή τη τελευταία περίπτωση ο ρόλος του Ηλεκτρονικού Υπολογιστή συνίσταται στη μετατροπή των δεδομένων σε

κάποιο ιστόγραμμα ή καμπύλη που να δίνει όσο γίνεται πιο περιεκτικά και παραστατικά τις απαραίτητες πληροφορίες στο διοικητικό στέλεχος. Το Πληροφοριακό Σύστημα Διοίκησης σε αυτή την περίπτωση χαρακτηρίζεται σαν **COMPUTER BASED INFORMATION SYSTEM**.

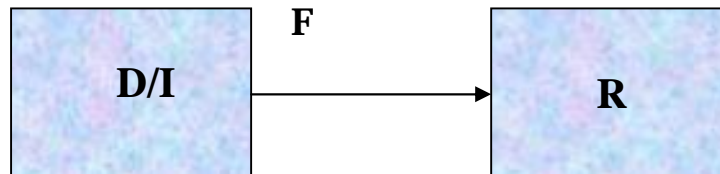
Οι πληροφορίες λοιπόν είναι επεξεργασμένα δεδομένα σε οργανωμένη και χρησιμοποιήσιμη μορφή και μπορούν να χωριστούν σε ποιοτικές ή ποσοτικές.



Σχήμα 1-5: Πληροφορίες

Πληροφορία επίσης χαρακτηρίζεται κάθε κριτήριο επιλογής μεταξύ των στοιχείων ενός συνόλου, δηλαδή, κάθε κριτήριο που επιτρέπει τον περιορισμό του μεγέθους αυτού του συνόλου σε ένα υποσύνολο που περιέχει την απάντηση μιας συγκεκριμένης ερώτησης (ελάττωση αβεβαιότητας). Αν για παράδειγμα, η ερώτηση είναι ποια είναι η μεγαλύτερη σε μήκος λέξη της ελληνικής γλώσσας με δεδομένες τις πληροφορίες ότι η λέξη που αναζητάμε έχει μήκος μεγαλύτερο από 5 γράμματα και είναι επίρρημα (κριτήρια), τότε το πρόβλημα περιορίζεται στην αναζήτηση της λύσης μέσα στο υποσύνολο των λέξεων της ελληνικής γλώσσας, που είναι επίρρηματα και έχουν μήκος μεγαλύτερο από 5 γράμματα. Είναι φανερό ότι στον παραπάνω ορισμό σαν πληροφορία θεωρείται κάθε χρήσιμο δεδομένο στην αναζήτηση της λύσης ενός προβλήματος, επομένως αυτή θεωρείται σαν μέρος των αρχικών δεδομένων

και όχι σαν αποτέλεσμα της επεξεργασίας αυτών. Θα πρέπει να τονιστεί ότι σύμφωνα με αυτή την αντίληψη αποκτά σημαντικό ενδιαφέρον η έννοια της επεξεργασίας της πληροφορίας η οποία μπορεί να οριστεί σαν μια συνάρτηση αναγωγής $f(D,I)$ των αρχικών δεδομένων D με βάση τις διαθέσιμες πληροφορίες I σε ένα σύνολο αποτελεσμάτων R (Σχήμα 1-6).



Σχήμα1-6: Επεξεργασία της Πληροφορίας

Οι πηγές των πληροφοριών μπορεί να είναι εσωτερικές (**internal**) ή εξωτερικές (**external**). Οι εσωτερικές πληροφορίες αντλούνται από διάφορες αναφορές, μελέτες και στατιστικά δεδομένα που αφορούν τα προϊόντα και τις υπηρεσίες που παρέχει η ίδια η επιχείρηση. Οι εξωτερικές πληροφορίες αντλούνται από κοινωνικούς, πολιτικούς, οικονομικούς και τεχνολογικούς παράγοντες. Η εξωτερική πληροφόρηση σε μία επιχείρηση χρησιμοποιείται από την διοίκηση στη δημιουργία του επιχειρησιακού σχεδίου(**business plan**).

1.2.2 ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΠΛΗΡΟΦΟΡΙΑΣ

Η **χρησιμότητα** και η **αξία** της **πληροφορίας** εξαρτάται από το βαθμό στον οποίον ικανοποιούνται τα παρακάτω χαρακτηριστικά της:

- **Σχετικότητα(Relevance)**
- **Πληρότητα(Completeness)**
- **Ακρίβεια(Accuracy)**
- **Επικαιρότητα(Timeliness)**
- **Αποτελεσματικότητα Κόστους(Cost effectiveness)**

ο Ελεγχιμότητα(Audibility)

Όταν η πληροφορία έχει τα παραπάνω χαρακτηριστικά θεωρείται αξιόπιστη(**Reliable**).

Για να προκύψει η αξία της πληροφορίας ,εξετάζουμε την εξίσωση [όφελος-κόστος=αξία]

Η επεξεργασία πληροφοριών είναι μια σημαντική λειτουργία στο χώρο της επιχείρησης με ευρύτερες κοινωνικές διαστάσεις.

Ένα μεγάλο μέρος της εργασίας και του προσωπικού χρόνου κάθε χρήστη καταναλώνεται στην καταγραφή, αναζήτηση και απορρόφηση πληροφοριών. Στατιστικές μελέτες έχουν δείξει ότι μέχρι το **80%** του χρόνου στελεχών επιχειρήσεων αναλύεται στην επεξεργασία και μεταφορά πληροφοριών.

Επίσης στις προηγμένες χώρες πλέον το **50%** του ανθρώπινου δυναμικού απασχολείται σε εργασίες που περιλαμβάνουν κάποια μορφή επεξεργασίας πληροφοριών-απλών κειμένων, αναφορών, αναλύσεων, σχεδίων δράσης κ.λ.π.

Οι Ηλεκτρονικοί Υπολογιστές έχουν εξελιχθεί σε αναπόσπαστο κομμάτι της οργανωμένης επεξεργασίας πληροφοριών λόγω των δυνατοτήτων τους και του μεγάλου όγκου δεδομένων που μπορούν να επεξεργαστούν.

Η χρήση των υπολογιστών στην επεξεργασία δεδομένων ξεκίνησε το **1954** όταν ένας από τους πρώτους υπολογιστές προγραμματίστηκε για να εκτελέσει λειτουργίες μισθοδοσίας. Σήμερα, η αυτοματοποιημένη διεκπεραίωση συναλλαγών σε μεγάλους οργανισμούς αποτελεί επιβεβλημένη λειτουργία. Ακόμη περισσότερο, η αυτοματοποιημένη επεξεργασία δεδομένων έχει επεκτείνει τις δυνατότητες εξαγωγής και χρήσης τυποποιημένων πληροφοριών. Η πρόκληση στην επεξεργασία δεδομένων σήμερα βρίσκεται στην εκμετάλλευση των δυνατοτήτων των υπολογιστών με σκοπό την υποστήριξη νοητικών εργασιών, όπως οι διοικητικές λειτουργίες και η λήψη αποφάσεων. Η μεγάλη ποικιλία των ηλεκτρονικών αυτών πόρων που χρησιμοποιούνται στην

διεκπαιρέωση συναλλαγών, στη παροχή επεξεργασμένων πληροφοριών σε τυποποιημένη μορφή, στην υποστήριξη διοικητικών λειτουργιών και στη λήψη αποφάσεων συγκροτούν τα Πληροφοριακά Συστήματα Διοίκησης.

Ο αυτοματισμός γραφείου, η τεχνολογία τηλεπικοινωνιών που υποστηρίζει την επικοινωνία μεταξύ ατόμων καθώς και τα μέσα υποστήριξης απλών λειτουργιών αποτελούν επίσης στοιχεία ενός ΠΣΔ. Ο σχεδιασμός και η υλοποίηση ενός ΠΣΔ σε έναν οργανισμό προϋποθέτει την αναγνώριση και καταγραφή των πληροφοριακών του απαιτήσεων. Οι απαιτήσεις για διεκπεραίωση συναλλαγών ρουτίνας είναι γενικά γνωστές και προσδιορίζονται εύκολα σε αντίθεση με τις πληροφοριακές απαιτήσεις της διοίκησης για την υποστήριξη αποφάσεων που εμφανίζουν μεγαλύτερη δυσκολία στο προσδιορισμό τους.

1.3 ΤΟ ΓΝΩΣΤΙΚΟ ΠΕΔΙΟ ΤΩΝ ΠΛΗΡΟΦΟΡΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ ΔΙΟΙΚΗΣΗΣ

Πολλά από τα στοιχεία που συγκροτούν την έννοια των ΠΣΔ μπορούν να απαντηθούν και σε άλλους επιστημονικούς κλάδους. Τέσσερις είναι οι κύριοι επιστημονικοί κλάδοι που έχουν άμεση σχέση και συμβάλλουν στην ανάπτυξη των ΠΣΔ: Οικονομικές Επιστήμες, Επιχειρησιακή Έρευνα, Θεωρία Διοίκησης και Οργάνωσης και Επιστήμη των Ηλεκτρονικών Υπολογιστών.

1.3.1 ΟΙΚΟΝΟΜΙΚΕΣ ΕΠΙΣΤΗΜΕΣ

Μεταξύ των Οικονομικών Επιστημών, ο Τομέας της Λογιστικής χαρακτηρίζεται από δύο περιοχές ενδιαφέροντος σε σχέση με τα ΠΣΔ: το καθαρά λογιστικό μέρος και το διαχειριστικό. Το καθαρά λογιστικό μέρος

ασχολείται με την καταχώρηση εσόδων-εξόδων για συγκεκριμένες χρονικές περιόδους, όπως κάθε μήνα ή έτος ή την έκδοση του ισολογισμού στο τέλος μιας περιόδου. Οι απολογισμοί τέτοιων περιόδων απευθύνονται προς πιθανούς επενδυτές και δημόσιους φορείς και κατά συνέπεια το καθαρά λογιστικό μέρος έχει περιορισμένη χρησιμότητα όσον αφορά λήψη διοικητικών αποφάσεων. Το διαχειριστικό όμως μέρος ενδιαφέρει για το προσδιορισμό συναφών δαπανών και την εκτέλεση αναλύσεων που χρησιμεύουν στον διαχειριστικό έλεγχο και τη λήψη αποφάσεων σε επίπεδο διοίκησης. Εστιάζει στην προετοιμασία προϋπολογισμών και την ανάλυση της απόδοσης σύμφωνα με τον προϋπολογισμό.

Η έννοια των ΠΣΔ είναι ασφαλώς πολύ ευρύτερη από ότι το διαχειριστικό μέρος ενός οικονομικού συστήματος. Τα Πληροφοριακά Συστήματα που παρέχουν στους χρήστες πρόσβαση σε δεδομένα και μοντέλα ,δεν ταυτίζονται με το κλασικό σύστημα λογιστικής διαχείρισης. Σήμερα η οργανωτική πρακτική περιορίζει την έκδοση και ανάλυση προϋπολογισμών στα λογιστικά τμήματα αφήνοντας το ΠΣΔ να προσφέρει την αναγκαία υποστήριξη σε δεδομένα και μοντέλα για παροχή των αναγκαίων πληροφοριών για Λειτουργικό Έλεγχο και Διοίκηση.

1.3.2 ΕΠΙΧΕΙΡΗΣΙΑΚΗ ΈΡΕΥΝΑ

Το αντικείμενο της Επιχειρησιακής Έρευνας συνίσταται στην εφαρμογή επιστημονικών μεθόδων και ποσοτικών αναλύσεων για επίλυση προβλημάτων απόφασης. Οι τεχνικές και μέθοδοι που χρησιμοποιούνται χαρακτηρίζονται από:

- 1.Έμφαση στη συστηματική προσέγγιση επίλυσης προβλημάτων.
- 2.Εφαρμογή μαθηματικών μοντέλων , καθώς και μαθηματικών και στατιστικών μεθόδων ανάλυσης.

3.Αναζήτηση βέλτιστης απόφασης ή βέλτιστης πολιτικής.

Η Επιχειρησιακή Έρευνα είναι σημαντική όσον αφορά τη σχέση της με τα ΠΣΔ γιατί διαμορφώνει και αναπτύσσει διαδικασίες ανάλυσης και υποστήριξης, με χρήση Ηλεκτρονικών Υπολογιστών ,πολλών τύπων προβλημάτων που σχετίζονται με τη λήψη αποφάσεων. Η συστηματική προσέγγιση στην επίλυση προβλημάτων, η χρησιμοποίηση μοντέλων και αλγόριθμων για λήψη αποφάσεων μέσω Η/Υ ενσωματώνονται σε ένα (υπο)Σύστημα Υποστήριξης Αποφάσεων (ΣΥΑ) ενός ΠΣΔ.

1.3.3 ΘΕΩΡΙΑ ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΟΡΓΑΝΩΣΗΣ

Εφόσον τα ΠΣΔ είναι συστήματα υποστήριξης και ελέγχου οργανωτικών λειτουργιών, επηρεάζονται άμεσα από πρακτικές οργάνωσης ,οργανωτικής συμπεριφοράς, διαχείρισης και λήψης αποφάσεων που εφαρμόζονται σε έναν Οργανισμό. Τα πεδία της διαχείρισης (ή της οργανωτικής συμπεριφοράς)και της οργανωτικής θεωρίας παρέχουν αρκετές έννοιες οι οποίες είναι κλειδιά για τη κατανόηση της λειτουργίας ενός ΠΣΔ σε έναν Οργανισμό. Μερικές από αυτές τις έννοιες είναι:

- 1.Θεωρία της οργανωτικής συμπεριφοράς και μεμονωμένης λήψης αποφάσεων
- 2.Μεμονωμένα κίνητρα συμπεριφοράς
- 3.Ομαδοποιημένες επεξεργασίες και λήψεις αποφάσεων
- 4.Τεχνικές ηγεσίας και διοίκησης
- 5.Διαδικασίες Οργανωτικών αλλαγών
- 6.Οργανωτική δομή και σχεδιασμός

1.3.4 ΕΠΙΣΤΗΜΗ ΤΩΝ ΗΛΕΚΤΡΟΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ ΚΑΙ ΠΛΗΡΟΦΟΡΙΚΗ

Η επιστήμη των Ηλεκτρονικών Υπολογιστών είναι σημαντική στο βαθμό που καλύπτει αντικείμενα με τα οποία σχετίζεται ο τρόπος επεξεργασίας δεδομένων σε ένα ΠΣΔ. Τέτοια αντικείμενα είναι η αλγοριθμική ,η υπολογιστική θεωρία ,το λογισμικό και οι δομές δεδομένων. Πάντως ,σε ακαδημαϊκό επίπεδο ,τα ΠΣΔ δεν αποτελούν τόσο προέκταση της επιστήμης των Ηλεκτρονικών Υπολογιστών, όσο της Θεωρίας Διοίκησης και Οργάνωσης. Οι θεμελιώδεις διαδικασίες σχεδιασμού και λειτουργίας ενός ΠΣΔ σχετίζονται και επηρεάζονται περισσότερο από τις διαδικασίες οργάνωσης και την οργανωτική απόδοση παρά από τους υπολογιστικούς αλγόριθμους. Η έμφαση των ΠΣΔ σε σχέση με την Πληροφορική δίνεται στην εφαρμογή των τεχνικών δυνατοτήτων που προσφέρει η επιστήμη των Ηλεκτρονικών Υπολογιστών.

ΚΕΦΑΛΑΙΟ 2^ο

ΓΕΝΙΚΗ ΘΕΩΡΗΣΗ ΚΑΙ ΚΑΤΑΤΑΞΗ ΠΛΗΡΟΦΟΡΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ ΔΙΟΙΚΗΣΗΣ

ΕΙΣΑΓΩΓΗ

Στο κεφάλαιο αυτό τεκμηριώνεται η αναγκαιότητα των Πληροφοριακών Συστημάτων Διοίκησης και περιγράφονται οι κατηγορίες και τα συστατικά μέρη των συστημάτων αυτών. Ο κύριος σκοπός του κεφαλαίου είναι η κατανόηση του σκοπού, των λειτουργιών και της γενικής δομής των Πληροφοριακών Συστημάτων Διοίκησης, ανάλογα με τις ιδιαίτερες ανάγκες των Διοικητικών μηχανισμών και τα επίπεδα της ιεραρχίας που απευθύνονται...

2.1. Η ΕΝΝΟΙΑ ΤΟΥ ΣΥΣΤΗΜΑΤΟΣ

Από αυτά που αναφέρθηκαν στο προηγούμενο κεφάλαιο , δυο πράγματα γίνονται σαφή:

α) Η έννοια του ΠΣΔ είναι τόσο παλιά όσο και η ύπαρξη οργανωμένων μονάδων για παραγωγή υπηρεσιών και προϊόντων.

β) Η επινόηση και αξιοποίηση των Η/Υ από τον άνθρωπο οδήγησε στην ανάπτυξη τεχνολογίας ΠΣΔ βασισμένη στον Η/Υ με βασικό χαρακτηριστικό την πολύ μεγάλη ταχύτητα επεξεργασίας.

Παρ' όλα αυτά το γεγονός ότι η καθιέρωση της χρήσης των Η/Υ έγινε πολύ γρήγορα και χωρίς ίσως να προηγηθεί μια ανάλογη εκπαίδευση σε όλες τις

βαθμίδες του προσωπικού –που έμμεσα ή άμεσα εξαρτώνται από τη λειτουργία του ΠΣΔ –είχε σαν αποτέλεσμα να αμφισβητείται συχνά η σκοπιμότητα εξέλιξης ενός χειροκίνητου σε ένα αυτοματοποιημένο, με τη χρήση Η/Υ, Πληροφοριακό Σύστημα Διοίκησης.

Στον **Πίνακα 2.1** φαίνονται οι πιο χαρακτηριστικές διαφορές των δυο κατηγοριών συστημάτων.

	<i>Χειρογραφικό Σύστημα</i>	<i>Σύστημα με Η/Υ</i>
1. Κατανόηση της Τεχνολογίας	Έυκολη. Συνήθως ανθρώπινη επεξεργασία η λειτουργία απλών μηχανών.	Δύσκολη. Πρωτότυπη και ελάχιστα κατανοητή τεχνολογία από το μέρος των χρηστών
2. Ανάπτυξη Προδιαγραφών	Πολύ περιγραφικές και εύκολα μεταβαλλόμενες.	Μια τυπική διαδικασία. Απαιτεί ακρίβεια και λεπτομέρεια
3. Διαχείριση Ανάπτυξης του συστήματος	Έυκολη με πολύ απλές διαδικασίες.	Πολύ δύσκολο να συμπληρωθεί χωρίς καθυστέρηση και χωρίς υπέρβαση προϋπολογισμού.
4. Μετατροπή και εγκατάσταση	Συνήθως εύκολη διαδικασία, με λίγες προσθήκες η αλλαγές.	Μπορεί να είναι μια δύσκολη διαδικασία που να απαιτεί αλλαγές και εκπαίδευση προσωπικού.
5. Εκαμψία	Συνήθως Μεγάλη. Ευκολία αλλαγών Διαδικασιών.	Συνήθως Μικρή. Συχνά πολύ δύσκολες οι μετατροπές ή αλλαγές διαδικασιών και με πολύ υψηλό κόστος
6. Απόδοση	Μικρής ταχύτητας και ακριβείας με περιορισμένη δυνατότητα επεξεργασίας	Μεγάλη ταχύτητα και ακρίβεια. Φθινή λειτουργία και δυνατότητα επεξεργασίας όγκου πληροφοριών

Πίνακας 2-1: Σύγκριση χειρογραφικών Πληροφοριακών Συστημάτων Διοίκησης και Πληροφορικών Συστημάτων Διοίκησης που στηρίζονται στον Η/Υ.

Από την σύγκριση αυτή βγαίνει ένα πολύ σημαντικό συμπέρασμα :ότι το μοναδικό χαρακτηριστικό υπεροχής των ΠΣΔ που στηρίζονται στους Ηλεκτρονικούς Υπολογιστές έναντι των χειροκίνητων είναι η απόδοση του Συστήματος που συχνά αποτελεί και καταλυτικό παράγοντα επιλογής της

μορφής του Συστήματος που θα εγκατασταθεί στην Επιχείρηση. Ένας Τραπεζικός Οργανισμός , μια Αεροπορική Εταιρεία ,μια αλυσίδα **SUPER MARKET** δεν είναι δυνατόν να ανταποκριθούν στις λειτουργικές απαιτήσεις χωρίς την υποστήριξη από ένα ΠΣΔ που στηρίζεται στους Ηλεκτρονικούς Υπολογιστές.

2.2. ΠΛΗΡΟΦΟΡΙΑΚΑ ΣΥΣΤΗΜΑΤΑ ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΟΡΓΑΝΩΤΙΚΗ ΔΟΜΗ ΕΠΙΧΕΙΡΗΣΗΣ.

2.2.1 Η ΦΥΣΙΟΓΝΩΜΙΑ ΤΩΝ ΠΛΗΡΟΦΟΡΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ ΔΙΟΙΚΗΣΗΣ.

Υπάρχει μεγάλη διαφορά απόψεων όσο αφορά τον ορισμό αλλά και τα ειδικότερα χαρακτηριστικά ενός Πληροφοριακού Συστήματος Διοίκησης και έτσι εμφανίζεται και διαφορετική ορολογία στην διεθνή Βιβλιογραφία:“ Σύστημα Επεξεργασίας Πληροφοριών ”, “Σύστημα Πληροφοριών και Λήψης Αποφάσεων”, “Πληροφοριακό Σύστημα Οργάνωσης” ή απλά “Πληροφοριακό Σύστημα”. Όλοι αυτοί οι όροι περιγράφουν ένα σύστημα το οποίο υποστηρίζει την λειτουργία ,διοίκηση και λήψη αποφάσεων ενός οργανισμού με χρήση Η/Υ.

Ένα Πληροφοριακό Σύστημα Διοίκησης μπορεί να περιγραφεί σαν μια πυραμίδα, κατ’ αντιστοιχία με αυτή του Οργανωτικού σχήματος ,όπου το κατώτερο επίπεδο της αποτελείται από πληροφορίες για διεκπεραίωση συναλλαγών, διερεύνηση καταστάσεων κ.λπ. Το επόμενο επίπεδο αποτελείται από πληροφορίες χρήσιμες στην καθημερινή λειτουργία του οργανισμού, το τρίτο επίπεδο από πληροφορίες υποστηρίξης τακτικού σχεδιασμού και λήψης

αποφάσεων για διοικητικό έλεγχο και το τελευταίο επίπεδο από πληροφορίες υποστήριξης στρατηγικού σχεδιασμού και διαμόρφωσης πολιτικής .Κάθε επίπεδο αυτής της πυραμίδας μπορεί να χρησιμοποιηθεί σαν δεδομένα πληροφορίες κατώτερων επιπέδων αυτούσια ή σαν προϊόν σύνθεσης αυτών.



Σχήμα 2-1: Οργανωτικό σχήμα Πυραμίδας Πληροφοριακού συστήματος

**ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΩΝ ΠΛΗΡΟΦΟΡΙΑΚΟΥ
ΣΥΣΤΗΜΑΤΟΣ ΔΙΟΙΚΗΣΗΣ ΜΕΓΑΛΟΥ ΟΡΓΑΝΙΣΜΟΥ.**

Από το προηγούμενο σχήμα φαίνεται ότι το ΠΣΔ ενός Οργανισμού εκτείνεται και διαμορφώνεται ανάλογα, σε όλα τα επίπεδα της Οργανωτικής Πυραμίδας. Ας υποθέσουμε ότι αναφερόμαστε σε μεγάλο Τραπεζικό Οργανισμό. Χαρακτηριστική λειτουργία επεξεργασίας συναλλαγών είναι η συναλλαγή που έχουν οι πελάτες στα ταμεία των Υποκαταστημάτων της Τράπεζας (Αναλήψεις, καταθέσεις...)

Οι παρεχόμενες πληροφορίες από το ΠΣΔ σε αυτό το επίπεδο **υπαγορεύουν διοικητικές πράξεις**, λίγο-πολύ υποχρεωτικές στο χρήστη. Ένα ΠΣΔ που έχει μόνο αυτά τα χαρακτηριστικά ονομάζεται **Σύστημα Επεξεργασίας Συναλλαγών (Transaction Processing System-TPS)**. Τα συστήματα επεξεργασίας συναλλαγών μπορούν επίσης να χρησιμοποιηθούν από διοικητικούς υπαλλήλους των ανώτερων επιπέδων ,κυρίως για να δημιουργήσουν διάφορες αναφορές.

Οι Διευθυντές και υποδιευθυντές των Υποκαταστημάτων αυτού του ίδιου Οργανισμού στο τέλος κάθε ημέρας ζητούν στοιχεία για την κίνηση του κάθε ταμείου (Όγκος Συναλλαγών, αριθμός εξυπηρετηθέντων πελατών,...)που θα τους χρησιμεύσουν για τον έλεγχο της απόδοσης των υπαλλήλων, τον ανασχεδιασμό του ωραρίου λειτουργίας του κάθε ταμείου, τον προγραμματισμό των αδειών. Ο τύπος αυτής της πληροφορίας **περισσότερο κατευθύνει παρά υπαγορεύει** στον αποφασίζοντα τις ενέργειες για έλεγχο και σχεδιασμό των εργασιών του Υποκαταστήματος και χαρακτηρίζεται σαν προϊόν ενός **Δομημένου Συστήματος Αποφάσεων (Structured Decision System)**.

Στην Κεντρική Υπηρεσία του ίδιου Οργανισμού, η Διεύθυνση Υποκαταστημάτων οφείλει να αποφασίσει τους τόπους και τον χρόνο εγκατάστασης των νέων Υποκαταστημάτων (τακτικό σχεδιασμό)στα πλαίσια της στρατηγικής της Διοίκησης. Μια τέτοια απόφαση μπορεί να υποστηριχτεί από ένα **Σύστημα Υποστήριξης Αποφάσεων (Decision Support System)**το

οποίο θα λαμβάνει υπ' όψιν αφενός τα διάφορα κριτήρια αξιολόγησης χώρων που προσφέρονται και αφετέρου την κρίση του αποφασίζοντα.

Τέλος, το Δ.Σ. της Τράπεζας χαράσσει την πολιτική του με βάση την εμπειρία των μελών ,αλλά και ενός συνόλου πληροφοριών που παρέχονται από το Πληροφοριακό Σύστημα σαν αποτέλεσμα επεξεργασίας στοιχείων από μοντέλα προγραμματισμού πρόβλεψης και εξομείωσης.

Οι κύριες λειτουργίες των διαφόρων Πληροφοριακών Συστημάτων Διοίκησης ,μπορούν να συνοψιστούν στον παρακάτω πίνακα.

TRANSACTION PROCESSING SYSTEMS
<ol style="list-style-type: none">1. Ταξινόμηση δεδομένων2. Ενημέρωση δεδομένων3. Επεξεργασία δεδομένων4. Έλεγχος δεδομένων5. Εκτυπώσεις δεδομένων
STRUCTURED DECISION SYSTEM
<ol style="list-style-type: none">1. Προγραμματισμός πόρων2. Αναλύσεις οικονομικές3. Αποτιμήσεις περιουσιακών στοιχείων4. Συσχετίσεις επιχειρησιακών μεγεθών με τα αντίστοιχα κλαδικά στοιχεία5. Προβλέψεις μέσω ανάλυσης δεδομένων6. Έλεγχος χρηματικής ροής7. Στατιστικές εκτυπώσεις
DECISION SUPPORT SYSTEM
<ol style="list-style-type: none">1.Επίλυση μη δομημένων προβλημάτων μέσω μοντέλων2.Δίνονται απαντήσεις σε ερωτήματα του τύπου:what if ?3.Στηρίζονται σε external data4.Η κρίση του manager είναι καθοριστική στη λύση των προβλημάτων.

Στη συνέχεια και με στόχο την κατανόηση των ιδιαίτερων χαρακτηριστικών, αλλά και των Υποσυστημάτων που συγκροτούν ένα ΠΣΔ, θα παρουσιαστεί ξεχωριστά κάθε μια από τις έννοιες κλειδιά: Σύστημα Χρήστη-Υπολογιστή, Ολοκληρωμένο Σύστημα, Βάση Δεδομένων, Μοντέλα Προγραμματισμού και Αποφάσεων. Ακόμη θα παρουσιαστεί η σχέση των ΠΣΔ με άλλες συναφείς έννοιες όπως Διαχείριση Πόρων Πληροφοριακών (**Information Resource Management-IRM**), Συστήματα Υποστήριξης Αποφάσεων (**Decision Support Systems-DSS**) και Επεξεργασία Δεδομένων (**Data processing-DP**).

2.2.2 ΣΥΣΤΗΜΑ ΧΡΗΣΤΗ-ΥΠΟΛΟΓΙΣΤΗ

Θεωρητικά ένα ΠΣΔ μπορεί να υπάρξει και χωρίς την παρουσία Η/Υ. Είναι όμως οι δυνατότητες των Η/Υ που καθιστούν εφικτή τη λειτουργία του ΠΣΔ. Το ερώτημα όμως δεν είναι αν θα πρέπει οι Η/Υ να χρησιμοποιούνται σε ένα ΠΣΔ αλλά σε ποια έκταση οι πληροφορίες και η διαχείρισή τους πρέπει να αυτοματοποιείται. Η έννοια του συστήματος χρήστη-υπολογιστή συνίσταται στο ότι ορισμένες εργασίες εκτελούνται από χρήστες ενώ άλλες εκτελούνται από μηχανές. Ο χρήστης ενός ΠΣΔ είναι οποιοσδήποτε άνθρωπος ασχολείται με την εισαγωγή δεδομένων ,παροχή οδηγιών προς το σύστημα ή με την αξιοποίηση των εξαχθέντων πληροφοριών. Ο χρήστης και ο υπολογιστής αποτελούν ένα συγκροτημένο σύνολο του οποίου τα αποτελέσματα προκύπτουν από τη συνεργασία μεταξύ τους.

Η συνεργασία χρήστη-υπολογιστή εξυπηρετείται λειτουργικά με τη σύνδεση της συσκευής εισαγωγής-εξαγωγής δεδομένων του χρήστη (συνήθως η οθόνη και το πληκτρολόγιο του τερματικού) με έναν υπολογιστή.

Ο υπολογιστής μπορεί να είναι ένας προσωπικός υπολογιστής που εξυπηρετεί μόνο ένα χρήστη (**single user**) ή ένας μεγάλος υπολογιστής ο οποίος εξυπηρετεί ένα μεγάλο αριθμό χρηστών μέσω των συνδεδεμένων σε αυτών

τερματικών. Η συσκευή εισαγωγής –εξαγωγής δεδομένων του χρήστη επιτρέπει την άμεση εισαγωγή δεδομένων και εξαγωγή αποτελεσμάτων. Για παράδειγμα , ένας χρήστης που ασχολείται με οικονομικό προγραμματισμό εισάγει μια ερώτηση τύπου “τι θα γίνει αν” χρησιμοποιώντας το πληκτρολόγιο του Η/Υ και λαμβάνει την απάντηση σε μερικά δευτερόλεπτα στην οθόνη.

Τα χαρακτηριστικά ενός αυτοματοποιημένου ΠΣΔ επηρεάζονται άμεσα από τις απαιτήσεις των χρηστών των τεχνολογικών δεδομένων και των γνώσεων των σχεδιαστών –κατασκευαστικών. «Αυτοματοποιημένο» σημαίνει ότι ο σχεδιαστής ενός ΠΣΔ πρέπει να έχει γνώση των Η/Υ και της χρησιμοποίησής τους στην επεξεργασία πληροφοριών. Η έννοια “χρήστης-υπολογιστής” σημαίνει ότι ο σχεδιαστής του συστήματος θα πρέπει να αντιλαμβάνεται τον χρήστη ως μέρος του συστήματος , καθώς και τη συμπεριφορά του ως τελικού παραλήπτη και χρήστη της πληροφορίας.

Οι εφαρμογές Πληροφοριακών Συστημάτων δεν θα έπρεπε να θεωρούν τους χρήστες ως επιστήμονες ειδικευμένους στην Πληροφορική. Από την άλλη πλευρά , οι χρήστες θα πρέπει να είναι σε θέση να προσδιορίσουν επακριβώς τις απαιτήσεις τους σε πληροφόρηση , να έχουν απλή γνώση των Η/Υ, καθώς και της φύσης και σιοπιμότητας των πληροφοριών στις διάφορες διοικητικές λειτουργίες.

2.2.3 ΟΛΟΚΛΗΡΩΜΕΝΟ ΣΥΣΤΗΜΑ

Η έννοια του **Ολοκληρωμένου Συστήματος** γίνεται περισσότερο κατανοητή μέσα από τα χαρακτηριστικά ενός μη Ολοκληρωμένου ΠΣΔ , το οποίο χαρακτηρίζεται έτσι όταν:

- ο Μεμονωμένες εφαρμογές στα πλαίσια ενός Πληροφοριακού Συστήματος αναπτύσσονται για διάφορους τύπους χρηστών, χωρίς να υπάρχει η αναγκαία διασύνδεση μεταξύ τους οπότε οι εφαρμογές χαρακτηρίζονται ασυνεπείς και ασύμβατες,

- Τα δεδομένα προσδιορίζονται ξεχωριστά για κάθε εφαρμογή και δεν μπορούν να χρησιμοποιηθούν από τις εφαρμογές στο σύνολό τους,
- Υπάρχουν διπλά αντίγραφα εφαρμογών που εξυπηρετούν τον ίδιο σκοπό,
- Ένας χρήστης, ο οποίος επιθυμεί να πραγματοποιήσει μια ανάλυση πληροφοριών ,χρησιμοποιεί δεδομένα από δυο ή περισσότερες διαφορετικές εφαρμογές.

Όταν συμβαίνει κάτι από τα παραπάνω, το ΠΣΔ δεν μπορεί να χαρακτηριστεί Ολοκληρωμένο.

Το πρώτο βήμα στην διασύνδεση εφαρμογών ανόμοιων πληροφοριακών υποσυστημάτων προς ένα Ολοκληρωμένο ΠΣΔ είναι ο καθολικός προγραμματισμός. Αν και οι εφαρμογές ενός Πληροφοριακού Συστήματος αναπτύσσονται ξεχωριστά , ο σχεδιασμός θα πρέπει να είναι καθολικός και να προσδιορίζει τον τρόπο διασύνδεσης των εφαρμογών. **Στην ουσία ένα ΠΣΔ σχεδιάζεται σαν ένα σύνολο μικρότερων πληροφοριακών συστημάτων.**

Ένα ΠΣΔ θα πρέπει να θεσπίσει πρότυπα ,γενικές οδηγίες και ρουτίνες τις οποίες θα πρέπει να ακολουθεί η ανάπτυξη κάθε εφαρμογής. Η συμμόρφωση ως προς αυτά επιτρέπει σε ανόμοιες εφαρμογές να μοιράζονται δεδομένα και να επιτρέπουν και αυτές με τη σειρά τους τη πρόσβαση σε διάφορους χρήστες.

Η τάση που υπάρχει σήμερα στο σχεδιασμό πληροφοριακών συστημάτων είναι η ύπαρξη μιας κεντρικής Βάσης Δεδομένων σαν ένα σύνολο διασυνδεδεμένων αρχείων που να επιτρέπει εξαγωγές πληροφοριών διαθέσιμων σε μια πλειάδα εφαρμογών και χρηστών.

2.2.4 Η ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ

Οι όροι "πληροφορίες" και "δεδομένα" χρησιμοποιούνται πολύ συχνά χωρίς να προσδιορίζεται σαφώς η διαφορά τους. "Πληροφορίες", όπως αναφέρεται και στο 1^ο κεφάλαιο, είναι δεδομένα τα οποία έχουν σημασία ή είναι χρήσιμα στο τελικό αποδέκτη τους ενώ "δεδομένα" είναι "πρώτη ύλη" από την οποία προκύπτει η "πληροφορία".

Η ύπαρξη της Βάσης Δεδομένων συνεπάγεται ότι τα δεδομένα πρέπει να είναι εκμεταλλεύσιμα και επεξεργάσιμα κατά τέτοιο τρόπο ώστε να εξάγονται πληροφορίες χρήσιμες για την Διοίκηση. Αυτή η διαδικασία προϋποθέτει οργάνωση των δεδομένων και επεξεργασία με την βοήθεια λογισμικού. Το λογισμικό που εκτελεί αυτή την εργασία ονομάζεται **Σύστημα Διαχείρισης Βάσεων Δεδομένων-ΣΔΒΔ(Data Base Management System-DBMS)**.

Όταν η πρόσβαση και η χρήση της Βάσης Δεδομένων ελέγχεται από ένα Σύστημα Διαχείρισης Βάσεων Δεδομένων, τότε όλες οι εφαρμογές που απαιτείται να έχουν πρόσβαση σε κάποια συγκεκριμένα δεδομένα, την έχουν χωρίς να απαιτείται αυτά να βρίσκονται αποθηκευμένα σε περισσότερα από ένα σημεία.. Ένα Ολοκληρωμένο Πληροφοριακό Σύστημα απαιτεί μια κεντρική διαχείριση της Βάσης Δεδομένων. Τα δεδομένα μπορεί να είναι αποθηκευμένα σε ένα κεντρικό υπολογιστή ή σε ένα αριθμό διασυνδεδεμένων υπολογιστών, αλλά με την προϋπόθεση ότι υπάρχει κάποια οργανωτική λειτουργία η οποία και διασφαλίζει την **ακεραιότητα (integrity)**, τον μη πλεονασμό(**non-redundancy**) και την **εξαντλητικότητα (exhaustivity)**.

2.2.5 ΧΡΗΣΙΜΟΠΟΙΗΣΗ ΜΟΝΤΕΛΩΝ

Συχνά οι ακατέργαστες ή συνοπτικές πληροφορίες δεν προσφέρουν σε ένα χρήστη αυτό που ζητάει. Συχνά απαιτείται τα δεδομένα να επεξεργάζονται και να παρουσιάζονται με τέτοιο τρόπο, ώστε να συμβάλλουν ουσιαστικά στη διαδικασία λήψης αποφάσεων. Για την υλοποίηση κάτι τέτοιου ,πρέπει να χρησιμοποιηθεί ένα πρότυπο (μοντέλο) αποφάσεων. Για παράδειγμα, μια απόφαση που αφορά στην επένδυση νέων κεφαλαίων μπορεί να ληφθεί βάσει ενός μοντέλου πρόβλεψης για επένδυση νέων κεφαλαίων.

Μοντέλα πρόβλεψης χρησιμοποιούνται σε διάφορα στάδια στην διαδικασία λήψης αποφάσεων. "Έξυπνα " μοντέλα μπορούν να χρησιμοποιηθούν στην αναζήτηση προβλημάτων ή επενδύσεων και στην αναγνώριση ή ανάλυση πιθανών λύσεων , ενώ άλλα μοντέλα μπορούν να χρησιμοποιηθούν στην ανεύρεση και βελτιστοποίηση της επιθυμητής λύσης.

Με άλλα λόγια , χρειάζονται πολλαπλές προσεγγίσεις για την υποστήριξη της διαδικασίας λήψης αποφάσεων. Τα ακόλουθα είναι ορισμένα παραδείγματα προβλημάτων και τύπων μοντέλων που μπορούν να περιλαμβάνονται σε ένα ΠΣΔ για την ανάλυση και υποστήριξη της λήψης αποφάσεων:

Πρόβλημα	Παράδειγμα Μοντέλου
Ποσότητα ασφαλούς αποθέματος προϊόντος	Μοντέλο το οποίο υπολογίζει το απόθεμα ασφαλείας και κάτω από υποθετικά σενάρια.
Επιλογή προσωπικού	Αναζήτηση προσωπικού και χρήση μοντέλου Πολυκριτήριας Ανάλυσης για ιεράρχηση εναλλακτικών λύσεων.
Κοστολόγηση νέων Προϊόντων	Μοντέλο εξομοίωσης ή πρόβλεψης εισαγωγής νέων προϊόντων
Έλεγχος εξόδων	Μοντέλο ελέγχου προϋπολογισμού

Σε ένα ευρύ Πληροφοριακό Σύστημα , ο υπεύθυνος για τη λήψη των αποφάσεων έχει στη διάθεσή του-εκτός από ένα σύνολο γενικών

μοντέλων(Στατιστικά Επιχειρησιακής Έρευνας κ.λπ.) τα οποία μπορούν να εφαρμοστούν σε πολλές αναλύσεις και αποφάσεις –και ένα σύνολο συγκεκριμένων μοντέλων για την υποστήριξη λήψης ειδικών αποφάσεων(Πολυκριτηριακά, σε περιβάλλον αβεβαιότητας κ.λπ.). Παρόμοια μοντέλα είναι διαθέσιμα για προγραμματισμό και έλεγχο. Το σύνολο αυτών των μοντέλων είναι η βάση των μοντέλων ενός ΠΣΔ.

Τα μοντέλα είναι πιο αποτελεσματικά όταν αυτοί που τα χρησιμοποιούν μπορούν, με διάλογο, να προγραμματίσουν ή να επαναλάβουν τη διαδικασία λήψης αποφάσεων κάτω από διαφορετικές συνθήκες **(trial-error process)**. Στην περίπτωση αυτή ονομάζονται **αλληλεπιδραστικά(interactive)**.

2.3. Η ΕΞΕΛΙΞΗ ΤΩΝ ΠΛΗΡΟΦΟΡΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ ΔΙΟΙΚΗΣΗΣ

Όταν πρωτοπαρουσιάστηκε η έννοια των ΠΣΔ, πολλοί οραματίστηκαν ένα μοναδικό ολοκληρωμένο σύστημα το οποίο μέσω των παραγόμενων πληροφοριών θα μπορούσε να απεικονίσει όλες τις απαραίτητες επεξεργασίες και λειτουργίες ενός οργανισμού. Άλλοι είχαν αμφιβολίες κατά πόσο θα μπορούσε να σχεδιαστεί ένα αυτοματοποιημένο πληροφοριακό σύστημα ,επαρκές να καλύψει τις ανάγκες που υπάρχουν στην υποστήριξη του προγραμματισμού και της λήψης αποφάσεων σε έναν οργανισμό και ειδικά του στρατηγικού σχεδιασμού. Ακόμη αμφιβολίες υπήρχαν όσον αφορά και την αξία χρησιμοποίησης προηγμένης τεχνολογίας , σε μια διαδικασία λήψης αποφάσεων η οποία δεν είναι σαφώς καθορισμένη.

Με την πάροδο του χρόνου ,η έννοια ενός μοναδικού ολοκληρωμένου συστήματος αποδείχτηκε ότι ήταν πολύπλοκη για να υλοποιηθεί. Η έννοια του ΠΣΔ εξελίχθηκε σε ένα σύνολο υποσυστημάτων τα οποία υλοποιούνται ,όπως και όταν χρειάζεται, αλλά ακολουθούν το γενικότερο σχεδιασμό ,πρότυπα και

ρουτίνες, όπως έχουν οριστεί για το ΠΣΔ. Έτσι αντί ενός μοναδικού, οικουμενικού ΠΣΔ ,ένας οργανισμός μπορεί να έχει πολλά διασυνδεδεμένα πληροφοριακά συστήματα που εξυπηρετούν τις ανάγκες της διοίκησης με ποικίλους τρόπους.

Τα ΠΣΔ ως έννοια συνεχίζουν να εξελίσσονται. Σχετίζονται, χωρίς να ταυτίζονται, με **Συστήματα Επεξεργασίας Δεδομένων** και άλλες έννοιες σχετικές με πληροφοριακά συστήματα. Δύο από αυτές τις έννοιες , οι οποίες θεωρούνται και προέκταση των ΠΣΔ ,είναι τα **Συστήματα Επιτελικής Πληροφόρησης και Συστήματα Διαχείρισης Πόρων Πληροφοριών**.

Μια πρόσφατα εμφανισθείσα τάση, που συμβαδίζει με την εξέλιξη των ΠΣΔ, είναι αυτή των **Συστημάτων Τελικού Χρήστη(end-user computing)**.

2.3.1 ΠΣΔ ΚΑΙ ΣΥΣΤΗΜΑΤΑ ΔΙΑΧΕΙΡΙΣΗΣ ΠΟΡΩΝ ΠΛΗΡΟΦΟΡΙΩΝ

Τα **Συστήματα Διαχείρισης Πόρων Πληροφοριών (ΣΔΠΠ)** είναι μια προσέγγιση στην υποστήριξη της διοίκησης ενός οργανισμού βασισμένη στην αντίληψη ότι οι πόροι πληροφοριών είναι και λειτουργικοί πόροι. Βάσει αυτής της αντίληψης ,η εργασία που πρέπει να εκτελέσει ένα πληροφορικό σύστημα είναι η σωστή διαχείριση αυτών των πόρων. Η έννοια "πόρος" είναι αρκετά ευρεία και μπορεί να περιλαμβάνει δεδομένα που λαμβάνονται μέσω εξοπλισμού επικοινωνιών , μέσω επεξεργασίας κειμένων, από προσωπικούς υπολογιστές καθώς και από παραδοσιακό τρόπο επεξεργασίας πληροφοριών. Τα δεδομένα μπορεί να προέρχονται από το εσωτερικό ή εξωτερικό περιβάλλον του Οργανισμού.

Τα τελευταία χρόνια η ανάπτυξη των **Δικτύων Πληροφοριών Ευρείας Ζώνης**- για τα οποία γίνεται ιδιαίτερη αναφορά στο Κεφάλαιο των Δικτύων –

οδήγησε στη δημιουργία Οργανισμών με αποκλειστικό σκοπό την διαχείριση και διάθεση πληροφοριών χρήσιμων για την λειτουργία των Επιχειρήσεων. Μεταξύ αυτών αναφέρεται το **Κέντρο Πληροφοριών της Ευρωπαϊκής Ένωσης ECHO (European Community Host Organisation)** με πληροφορίες για Επιχειρήσεις διαφόρων κλάδων της ΕΕ, τιμές προϊόντων, κλεισιμο χρηματιστηρίων κ.λπ. Σε εθνικό επίπεδο επίσης κυριαρχεί μια τάση ανάπτυξης τέτοιων Συστημάτων ,όχι τόσο σε επίπεδο μεμονωμένης επιχείρησης ,άλλωστε το μέγεθος των Ελληνικών Επιχειρήσεων δεν προσφέρεται για κάτι ανάλογο-όσο σε επίπεδο συλλογικών φορέων(Γ.Σ.Ε.Ε., Σ.Ε.Β., Επιμελητήρια ,...)οι οποίοι αναλαμβάνουν για λογαριασμό των μελών τους (Ομοσπονδίες ,Επιχειρήσεις, Επαγγελματικές Ενώσεις κ.λπ.)την ανάπτυξη, λειτουργία και εκμετάλλευση Συστημάτων Διαχείρισης Πόρων Πληροφοριών. Μια τέτοια πρωτοβουλία ενισχύθηκε πρόσφατα από την Ευρωπαϊκή Ένωση στα πλαίσια του Κοινοτικού Προγράμματος **TELEMATIQUE**.

2.3.2 ΠΛΗΡΟΦΟΡΙΑΚΟ ΣΥΣΤΗΜΑ ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΤΕΛΙΚΟΙ ΧΡΗΣΤΕΣ

Μια πρόσφατα εμφανισθείσα τάση είναι αυτή του "ισχυρού" τελικού χρήστη ενός συστήματος (**end-user computing**). Οι χρήστες προμηθεύονται τερματικά ή προσωπικούς υπολογιστές υποστηριζόμενα από λογισμικό ποικίλων δυνατοτήτων(**front-end tools**), για πρόσβαση σε δεδομένα, ανάπτυξη μοντέλων και άμεση επεξεργασία πληροφοριών. Αυτή η εξέλιξη πραγματοποιείται και ενισχύεται από τις συνεχώς αυξανόμενες δυνατότητες και την πτώση του κόστους της τεχνολογίας. Ο τελικός χρήστης σήμερα είναι ένας επιστήμονας, τεχνολόγος ή και απλός υπάλληλος Διοίκησης "θωρακισμένος" όμως κατάλληλα ούτως ώστε με ελάχιστες γνώσεις Πληροφορικής να μπορεί να επιτύχει άριστη εκμετάλλευση

ενός ΠΣΔ. Το παραθυρικό περιβάλλον (**Windows**), το οποίο έχει επικρατήσει σήμερα σαν πλατφόρμα ανάπτυξης και χρήσης εφαρμογών ,αποτελεί μορφή τέτοιων εργαλείων.

2.3.3 ΣΥΣΤΗΜΑΤΑ ΕΠΙΤΕΛΙΚΗΣ ΠΛΗΡΟΦΟΡΗΣΗΣ

Ο σκοπός ενός Συστήματος Επιτελικής Πληροφόρησης (ΣΕΠ) είναι να εξυπηρετεί τις πληροφοριακές ανάγκες των ανώτερων βαθμίδων της ιεραρχίας. Στόχος του Συστήματος Επιτελικής Πληροφόρησης είναι να δίνει γρήγορη και εύκολη πρόσβαση σε λεπτομερείς πληροφορίες που παρουσιάζονται με απλή και κατανοητή μορφή (διαγράμματα, πίνακες, κείμενο).

Αυτό επιτυγχάνεται με τη χρήση φίλτρων για την εμφάνιση μόνο της χρήσιμης κάθε φορά πληροφορίας και με την υιοθέτηση εξαιρετικά φιλικών διεπαφών χρήσης που προσαρμόζονται στις ανάγκες των αποφασίζόντων.

Η πιο σημαντική δυνατότητα που υποστηρίζει ένα **Σύστημα Επιτελικής Πληροφόρησης** είναι η δυνατότητα πλοήγησης στα επίπεδα λεπτομέρειας της πληροφορίας (**drill down, roll up**) με βάση τις παραμέτρους που θέτει ο αποφασίζων. Με τη δυνατότητα αυτή, ο αποφασίζων αποκτά συνολική επιχειρηματική προοπτική των λειτουργιών που τον αφορούν, έχει τη δυνατότητα να παρακολουθεί τις διαδικασίες αυτές αποτελεσματικά σε κάθε επίπεδο ιεραρχίας του οργανισμού και τέλος να αναγνωρίζει τα προβλήματα ,μόλις αυτά εκδηλωθούν. Έτσι παρέχεται και ένα εργαλείο για την ταχύτητα ανίχνευση και ανάλυση επιχειρησιακών προβλημάτων και την υποστήριξη της έγκαιρης επέμβασης για την αντιμετώπισή τους. Η πληροφόρηση φτάνει στον αποφασίζοντα με τη μορφή αναφορών, όπου εμφανίζεται η πληροφορία που τον ενδιαφέρει ,με βάση των σχεδιασμό των μορίων πολυδιάστατων δεδομένων (**data marts**) της αποθήκης δεδομένων.

Στις αναφορές αυτές η πληροφορία εμφανίζεται σε φιλική προς το χρήστη μορφή και υποστηρίζεται η δυνατότητα πλοήγησης προς περισσότερο εξειδικευμένα ή πιο αφηρημένα επίπεδα πληροφορίας.

Τα τελευταία χρόνια εμφανίζεται μια τάση ολοκλήρωσης των διαφόρων τύπων ΠΣΔ με την ανάδειξη νέων «συνεργαζόμενων υποσυστημάτων» πάνω σε σύγχρονες τηλεπικοινωνιακές πλατφόρμες. Συστήματα Υποστήριξης Ατομικών και Ομαδικών Αποφάσεων ,Διαχείρισης Επιχειρησιακών Πόρων, Διαχείρισης Πελατειακών Σχέσεων, Ηλεκτρονικού Επιχειρείν και Εφοδιαστικής είναι τα πλέον σημαντικά και συνήθη στους σύγχρονους Οργανισμούς, για τα οποία γίνεται ειδική αναφορά στο 8^ο Κεφάλαιο.








2.4. ΥΠΟΣΥΣΤΗΜΑΤΑ ΕΝΟΣ ΠΛΗΡΟΦΟΡΙΑΚΟΥ ΣΥΣΤΗΜΑΤΟΣ ΔΙΟΙΚΗΣΗΣ

Τα ΠΣΔ παρουσιάζονται σαν μια αρκετά ευρεία οντότητα η οποία αποτελείται από έναν αριθμό υποσυστημάτων. Υπάρχουν δυο προσεγγίσεις στον διαχωρισμό ενός ΠΣΔ σε υποσυστήματα:

- σύμφωνα με τις λειτουργικές διαδικασίες.
- σύμφωνα με τις διοικητικές διαδικασίες.

2.4.1 ΥΠΟΣΥΣΤΗΜΑΤΑ ΛΕΙΤΟΥΡΓΙΩΝ

Δεδομένου ότι οι οργανωτικές λειτουργίες είναι κατά κάποιο τρόπο διαχωρισμένες όσον αφορά τις ενέργειες που περιλαμβάνουν και ορίζονται διοικητικά σαν ξεχωριστές αρμοδιότητες , ένα ΠΣΔ μπορεί να εννοηθεί ότι περιλαμβάνει έναν αριθμό πληροφοριακών υποσυστημάτων κατά λειτουργία. Μπορεί να υπάρχουν διαδικασίες και **modules** κοινά σε περισσότερα του ενός υποσυστήματα αλλά κάθε υποσύστημα λειτουργιών είναι μοναδικό όσον αφορά τις διαδικασίες ,τα προγράμματα, τα μοντέλα κ.λπ. που χρησιμοποιεί. Τυπικά κύρια υποσυστήματα λειτουργιών ενός επιχειρησιακού παραγωγικού οργανισμού είναι:

ΚΥΡΙΑ ΥΠΟΣΥΣΤΗΜΑΤΑ ΛΕΙΤΟΥΡΓΙΩΝ	ΜΕΡΙΚΕΣ ΤΥΠΙΚΕΣ ΧΡΗΣΕΙΣ
 Marketing	Προεπίτμηση και στρατηγική πωλήσεων ,ανάλυση πελατών και πωλήσεων.
 Παραγωγή	Σχεδιασμός και προγραμματισμός της παραγωγής ,ανάλυση ελέγχου κόστους.
 Διοικητική Μέριμνα	Προγραμματισμός απαιτήσεων σε προσωπικό, ανάλυση αποδοτικότητας, διαχείριση μισθών.
 Προσωπικό	Προγραμματισμός απαιτήσεων σε προσωπικό ανάλυση αποδοτικότητας, διαχείριση μισθών.
 Οικονομικά/Λογιστικά	Οικονομική ανάλυση ,κοστολόγηση, προσδιορισμός απαιτήσεων σε κεφάλαιο, έσοδα.
 Επεξεργασία Πληροφοριών	Σχεδιασμός Πληροφοριακών Συστημάτων, ανάλυση κόστους ωφέλειας.
 Διοίκηση	Στρατηγικός Σχεδιασμός ,καταμερισμός κονδυλίων.

2.4.2 ΥΠΟΣΥΣΤΗΜΑΤΑ ΟΡΓΑΝΩΤΙΚΩΝ ΕΠΙΠΕΔΩΝ

Μια άλλη προσέγγιση στην κατανόηση της δομής ενός πληροφοριακού συστήματος είναι σε σχέση με τα οργανωτικά επίπεδα τα οποία υποστηρίζει. Συνήθως τα υποσυστήματα διαδικασιών είναι χρήσιμα για περισσότερα του ενός υποσυστήματα οργανωτικών επιπέδων.

Παραδείγματα κύριων Υποσυστημάτων Διαδικασιών είναι:

ΥΠΟΣΥΣΤΗΜΑ ΟΡΓΑΝΩ- ΤΙΚΩΝ ΕΠΙΠΕΔΩΝ	ΜΕΡΙΚΕΣ ΤΥΠΙΚΕΣ ΧΡΗΣΕΙΣ
✓ Επεξεργασία Συναλλαγών	Επεξεργασία παραγγελιών, αποστολών και παραλαβών
✓ Λειτουργικός Έλεγχος	Προγραμματισμός διαδικασιών και Εκθέσεις αποδοτικότητας
✓ Διοικητικός Έλεγχος	Τυποποίηση προϋπολογισμών και καταμερισμού κονδυλίων
✓ Στρατηγικός Σχεδιασμός	Τυποποίηση στόχων και στρατηγικών σχεδίων.

Παρατηρείται ότι τα υποσυστήματα οργανωτικών επιπέδων αυτά ανταποκρίνονται στα επίπεδα της ιεραρχικής πυραμίδας που προσδιορίζει ένα ΠΣΔ(Σχήμα 2.1). Οι σχέσεις των υποσυστημάτων οργανωτικών επιπέδων σε σχέση με τα υποσυστήματα λειτουργιών παρουσιάζεται στο επόμενο σχήμα. Το σχήμα αυτό αποτελεί μια προέκταση και γενίκευση του προηγούμενου σχήματος προς την κατεύθυνση ενός **Ολοκληρωμένου Πληροφοριακού Συστήματος Διοίκησης**. Κάθε σκιασμένο τετραγωνίδιο αποτελεί ένα υποσύστημα υποστήριξης Διοικητικών πράξεων και αποφάσεων συγκεκριμένου λειτουργικού τομέα και διαδικασιών του χώρου της επιχείρησης.

Η έννοια της υποστήριξης αποφάσεων που αναφέρεται κυρίως σε διαδικασίες των υψηλότερων κλιμακίων της Οργανωτικής Πυραμίδας(Στρατηγικός

Σχεδιασμός, Διοικητικός Έλεγχος) προϋποθέτει την ύπαρξη μιας **Βάσης Μοντέλων Απόφασης**, που αντλούνται από τον χώρο των Εφαρμοσμένων Μαθηματικών (Επιχειρησιακή Έρευνα , Στατιστική,...) και ενός αντίστοιχου Συστήματος Διαχείρισης.

2.5. ΦΥΣΙΚΗ ΔΟΜΗ ΕΝΟΣ ΠΛΗΡΟΦΟΡΙΑΚΟΥ ΣΥΣΤΗΜΑΤΟΣ ΔΙΟΙΚΗΣΗΣ

Τα τμήματα του συνιστούν την φυσική δομή ενός ΠΣΔ είναι: ο εξοπλισμός(**Hardware**), το λογισμικό(**Software**), η Βάση Δεδομένων(**Database**), οι Διαδικασίες(**Procedures**) και το Προσωπικό Λειτουργίας(**Operations Personnel**). Στον πίνακα που ακολουθεί επεξηγούνται συνοπτικά οι παραπάνω όροι.

ΤΜΗΜΑ ΦΥΣΙΚΗΣ ΔΟΜΗΣ ΠΣΔ	ΠΕΡΙΓΡΑΦΗ
Εξοπλισμός	Αναφέρεται στον εξοπλισμό Πληροφορικής και στις υπόλοιπες διασυνδεδεμένες συσκευές. Σκοπός του εξοπλισμού είναι να εξυπηρετήσει τις εξής λειτουργίες: <ol style="list-style-type: none"> 1.Εισαγωγή Στοιχείων 2.Εξαγωγή Πληροφοριών 3.Δευτερεύουσα αποθήκευση για δεδομένα και Προγράμματα 4.Κεντρική Επεξεργασία (υπολογισμούς, ελέγχους και πρωτεύουσα μνήμη) 5.Επικοινωνίες
Λογισμικό	Το λογισμικό αποτελεί ευρεία έννοια που Χρησιμοποιείται να δηλώσει τις εντολές που Κατευθύνουν την λειτουργία του εξοπλισμού. Οι δυο βασικοί τύποι λογισμικού είναι το Λογισμικό Συστήματος (System Software) και το Λογισμικό Εφαρμογών (Application Software)
Βάση Δεδομένων	Η Βάση Δεδομένων περιλαμβάνει όλα τα δεδομένα με την απαραίτητη διασύνδεση που διαχειρίζεται το Λογισμικό Εφαρμογών. Ένα αυτόνομο τμήμα αποθηκευμένων δεδομένων συνήθως αποκαλείται αρχείο. Όλα τα δεδομένα αποθηκεύονται σε ταινίες, δίσκους κ.λπ. οι οποίες είναι μονάδες βοηθητικής ή περιφερειακής μνήμης.

Διαδικασίες	<p>Οι διαδικασίες αποτελούν φυσικό τμήμα της Δομής ενός ΠΣΔ καθώς υφίστανται κατεγραμμένες και χρησιμοποιούνται για την υποστήριξη της λειτουργίας του Συστήματος. Υπάρχουν οι παρακάτω βασικές κατηγορίες διαδικασιών:</p> <ol style="list-style-type: none"> 1.Χρήσης και Λειτουργίας 2.Επεξεργασίας 3.Ασφάλειας 4.Αναβάθμισης και Ανανέωσης
Προσωπικό Λειτουργίας	<p>Διαχειριστές Συστημάτων, Αναλυτές Συστημάτων, Προγραμματιστές, Προσωπικό Εισαγωγής Στοιχείων.</p>

ΚΕΦΑΛΑΙΟ 3ο

ΟΡΓΑΝΩΣΗ ΔΕΔΟΜΕΝΩΝ

ΕΙΣΑΓΩΓΗ

Στο κεφάλαιο αυτό παρουσιάζονται οι μέθοδοι οργάνωσης δεδομένων για αποτελεσματική αποθήκευση και επεξεργασία. Ο βασικός σκοπός του κεφαλαίου είναι να κατανοήσει ο αναγνώστης τις βασικές δομές και τρόπους αποθήκευσης των δεδομένων σε αρχεία και της οργάνωσής τους ώστε να διευκολύνεται η εξαγωγή των πληροφοριών που αξιοποιούνται από ένα ΠΣΔ . Ειδική αναφορά γίνεται στις βάσεις Δεδομένων.

3.1. ΣΥΜΒΑΤΙΚΕΣ ΜΕΘΟΔΟΙ ΟΡΓΑΝΩΣΗΣ ΔΕΔΟΜΕΝΩΝ

3.1.1 ΒΑΣΙΚΕΣ ΈΝΝΟΙΕΣ

Όπως αναφέρθηκε προηγουμένα ,η σύγχρονη επιχείρηση κατακλύζεται από ένα τεράστιο όγκο δεδομένων που προέρχονται από μια μεγάλη ποικιλία πηγών. Τα δεδομένα αυτά επεξεργάζεται με την βοήθεια ενός ΠΣΔ που έχει αναπτύξει και παράγει πληροφορίες οι οποίες υποβοηθούν στην ορθή λήψη αποφάσεων.

Τα δεδομένα υπονοούν ένα σύνολο στοιχείων, τα οποία οργανώνονται κατά τρόπο που να επιτρέπει τον προσδιορισμό τους, την αποθήκευσή τους , την ανάληψή τους-όταν χρειαστούν-και την επεξεργασία τους προκειμένου να παραχθούν πληροφορίες.

Τα βασικά συστατικά της οργανωτικής δομής επιχειρησιακών δεδομένων είναι:

α) Χαρακτήρας (CHARACTER)

Είναι η ελάχιστη μονάδα δεδομένων που μπορεί να αποθηκευτεί. Ο χαρακτήρας μπορεί να είναι ένα γράμμα ,όπως το Α ,ένα ψηφίο του δεκαδικού συστήματος ,όπως το 8 , όπως το (+), το (-), το (/) κ.λπ.

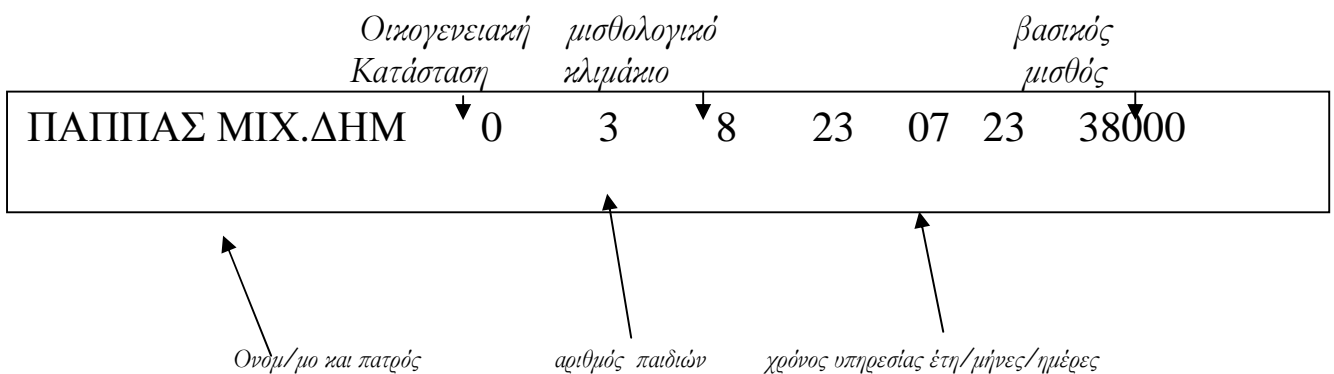
β)Πεδίο(FIELD)

Είναι μια ομάδα διαδοχικών χαρακτήρων που θεωρούνται ως ένα όλο και έχουν μια σημασία. Για παράδειγμα, το επώνυμο ενός πελάτη ή το ποσό που οφείλει ο πελάτης είναι πεδία.

Τα πεδία διακρίνονται σε "αριθμητικά" (NUMERIC) όταν αντιπροσωπεύουν αριθμητικές ποσότητες που συμμετέχουν σε αριθμητικές πράξεις και σε πεδία "χαρακτήρων"(STRINGS) τα οποία είναι απλώς ακολουθίες χαρακτήρων που αποθηκεύονται μόνον. Η ποσότητα που υπάρχει στην αποθήκη από ένα συγκεκριμένο είδος αποθέματος είναι αριθμητικό πεδίο , ενώ το επώνυμο ενός μισθοδοτημένου υπαλλήλου και ο αριθμός τηλεφώνου του είναι πεδία χαρακτήρων.

γ)Λογική εγγραφή(LOGICAL RECORD)

Είναι το σύνολο των πεδίων που περιλαμβάνει όλα τα λογικά συσχετισμένα στοιχεία για μια συγκεκριμένη επεξεργασία. Για παράδειγμα, όλα τα στοιχεία που απαιτούνται για να εξαχθεί η μισθοδοσία ενός υπαλλήλου αποτελούν μια λογική εγγραφή, μερικά από τα πεδία της οποίας φαίνονται στο Σχήμα 4.4.



Σχήμα3-1:Λογική Εγγραφή για τη Μισθοδοσία Υπαλλήλων.

δ) Αρχείο (FILE)

Είναι το σύνολο των λογικών εγγραφών που απαιτούνται για μια συγκεκριμένη επεξεργασία. Για παράδειγμα, το σύνολο όλων των λογικών εγγραφών για την μισθοδοσία των υπαλλήλων ενός οργανισμού συνιστά το αρχείο μισθοδοσίας. Τα αρχεία διακρίνονται ανάλογα με την χρήση τους στους επόμενους βασικούς τύπους:

"Κύριο αρχείο" (MASTER FILE)

Χρησιμοποιείται ως μόνιμη πηγή άντλησης πληροφοριών από τις μόνιμες λογικές εγγραφές που περιέχει. Για παράδειγμα, το κύριο αρχείο της αποθήκης συνίσταται από λογικές εγγραφές για κάθε αποθηκευμένο υλικό οι οποίες περιλαμβάνουν πεδία ταυτότητας υλικού, ποσότητα, τιμή, ιστορικές πληροφορίες διακίνησης υλικού κ.λπ.

"Αρχείο μεταβολών" (TRANSACTION FILE)

Περιλαμβάνει τις λογικές εγγραφές των πράξεων που πραγματοποιήθηκαν σε μια ορισμένη χρονική περίοδο, π.χ. σε μια ημέρα ή ένα μήνα. Χρησιμοποιείται κυρίως για ενημέρωση του κύριου αρχείου. Για παράδειγμα, το αρχείο τιμολογίων πωλήσεων, το αρχείο αγορών κ.λπ.

"Απολογιστικό αρχείο" (REPORT FILE)

Περιλαμβάνει λογικές εγγραφές οι οποίες περιέχουν στοιχεία από το κύριο αρχείο και χρησιμοποιείται για την σύνταξη απολογιστικών εκθέσεων. Για παράδειγμα, το αρχείο των κρατήσεων φόρου μισθωτών υπηρεσιών που δημιουργείται από το μισθολογικό μητρώο των υπαλλήλων ενός οργανισμού.

"Ιστορικό Αρχείο" (Historical File)

Περιλαμβάνει λογικές εγγραφές με στοιχεία που δεν βρίσκονται σε ισχύ σήμερα, αλλά θεωρούνται χρήσιμα κάτω από ορισμένες περιστάσεις, π.χ. το

Ιστορικό Αρχείο των υπαλλήλων ενός Οργανισμού περιλαμβάνει λογικές εγγραφές

με στοιχεία υπαλλήλων που αποχώρισαν για διάφορους λόγους από την Υπηρεσία.

ε) Βάση Δεδομένων (DATA BASE)

Είναι ένα σύνολο από αρχεία συσχετισμένα μεταξύ τους που συνιστούν μια οργανωμένη πηγή πληροφόρησης. Ενώ, κάθε απλό αρχείο προορίζεται, όπως προαναφέρθηκε, για μια συγκεκριμένη χρήση, όπως π.χ. το αρχείο μισθοδοσίας ή το αρχείο προσωπικού, η βάση δεδομένων μπορεί να χρησιμοποιηθεί για πολλές χρήσεις. Έτσι, αν το αρχείο μισθοδοσίας και το αρχείο διοικητικής κατάστασης προσωπικού, συσχετιστούν σε μια βάση δεδομένων, τότε η βάση αυτή θα μπορούσε να εξυπηρετήσει την μισθοδοσία, την παρακολούθηση της παραγωγικότητας προσωπικού κ.λπ.

Τα αρχεία αποθηκεύονται σε μέσα βοηθητικής μνήμης, κυρίως μαγνητικούς δίσκους. Λόγω της μεγάλης χωρητικότητας των μέσων αυτών, αρχεία που θα απαιτούσαν ένα τεράστιο όγκο καρτελών, βιβλίων και παραστατικών καταχωρούνται σε μερικούς μαγνητικούς δίσκους. Σε σπάνιες περιπτώσεις το αρχείο είναι τόσο μεγάλο, ώστε να απαιτήσει συστοιχία δίσκων, ενώ σε πολλές περιπτώσεις στον ίδιο δίσκο χωρούν περισσότερα του ενός αρχεία.

Έτσι τα μέσα αυτά μπορούν να φυλαχθούν **OFF-LINE** και όταν απαιτηθεί επεξεργασία των αρχείων πάνω σε αυτά να τοποθετηθούν στον μηχανισμό δίσκου για την επεξεργασία.

3.1.2 Τύποι Λογικών Εγγραφών

Ανάλογα με τις απαιτήσεις επεξεργασίας των αρχείων οι λογικές τους εγγραφές δομούνται κατάλληλα. Βασικά οι λογικές εγγραφές διακρίνονται σε δύο τύπους :τις λογικές εγγραφές "**σταθερού μήκους**" (**FIXED LENGTH RECORDS**) και τις λογικές εγγραφές "**μεταβλητού μήκους**" (**VARIABLE LENGTH RECORDS**).

α) Λογικές εγγραφές σταθερού μήκους

Με την τεχνική αυτής της δομής είναι γνωστά εκ των προτέρων – και αμετάβλητα για όλες τις λογικές εγγραφές – αφενός το μήκος σε χαρακτήρες κάθε πεδίου της λογικής εγγραφής, αφετέρου το πλήθος των πεδίων κάθε λογικής εγγραφής.

Κατά τον σχεδιασμό λοιπόν της λογικής εγγραφής σταθερού μήκους, προδιαγράφουμε εκ των προτέρων τα δεδομένα που θα περιληφθούν, τα πεδία που θα απαιτηθούν και το μήκος του κάθε πεδίου. Δεν υπάρχει περίπτωση να βρεθεί μια λογική εγγραφή σταθερού μήκους ένα πεδίο που δεν υπάρχει σε μια άλλη.

β) Λογικές εγγραφές μεταβλητού μήκους

Συχνά οι λογικές εγγραφές σταθερού μήκους οδηγούν σε σπατάλη χώρου του μέσου βοηθητικής μνήμης επειδή πολλά πεδία μπορεί να μένουν κενά. Τέτοια είναι η περίπτωση του αρχείου ασθενών ενός μικροβιολογικού εργαστηρίου το οποίο εκτελεί ένα ορισμένο αριθμό εξετάσεων. Εάν η λογική εγγραφή περιλαμβάνει πεδία δεδομένου μήκους για κάθε εξέταση είναι φανερό ότι αυτή σπάνια θα συμπληρωθεί, αφού κάθε ασθενής κάνει επιλεκτικά ένα μέρος μόνο των εξετάσεων. Με την τεχνική αυτή το πλήθος των πεδίων της λογικής

εγγραφής και το μήκος τους δεν είναι δεδομένα, αλλά μεταβάλλονται από λογική εγγραφή σε λογική εγγραφή.

3.1.3 Σειριακή Οργάνωση Αρχείων

Κατά την σειριακή οργάνωση των αρχείων, οι λογικές εγγραφές καταχωρούνται σειριακά, δηλαδή η μία μετά την άλλη, σύμφωνα με ένα **κλειδί (KEY)**. Σαν κλειδί μπορεί να χρησιμοποιηθεί ένα οποιαδήποτε πεδίο των λογικών εγγραφών του αρχείου, το ίδιο για όλες τις εγγραφές. Για παράδειγμα, το αρχείο των πελατών μιας επιχείρησης μπορεί να είναι καταχωρημένο στο μέσο βοηθητικής μνήμης κατά αλφαβητική σειρά του ονόματος ή κατά αύξουσα σειρά του κωδικού των πελατών.

Ανάληψη Πληροφοριών από σειριακό αρχείο

Υπάρχουν τρεις τύποι ανάληψης πληροφοριών από ένα σειριακό αρχείο:

α) **Επιλεκτική ανάληψη**, όπως για παράδειγμα η εκτύπωση ενός πίνακα των υπαλλήλων μιας επιχείρησης με ηλικία άνω των **40** ετών και προϋπηρεσία τουλάχιστον **6** έτη. Η επιλεκτική αυτή ανάληψη πληροφοριών από το σειριακό αρχείο του προσωπικού της επιχείρησης απαιτεί έρευνα όλων των λογικών εγγραφών του αρχείου μία προς μία. Σε κάθε λογική εγγραφή ελέγχονται τα πεδία "ηλικία υπαλλήλου" και "χρόνος προϋπηρεσίας" και εφόσον συμφωνούν με τις τεθείσες απαιτήσεις τότε το όνομα, η ηλικία και η προϋπηρεσία του υπαλλήλου εκτυπώνονται. Η διαδικασία επιλεκτικής ανάληψης που αναφέρθηκε ισχύει και για τις δύο περιπτώσεις που το σειριακό αρχείο βρίσκεται σε σειριακό μέσο βοηθητικής

μνήμης (π.χ μαγνητοταινία) ή βρίσκεται σε μέσο απευθείας προσπέλασης (π.χ μαγνητικός δίσκος).

β) **Εξαγωγική ανάληψη με κλειδί διαφορετικό** από αυτό με το οποίο έχει ταξινομηθεί το αρχείο. Για παράδειγμα, έστω ότι το αρχείο υπαλλήλων είναι ταξινομημένο κατά τον κωδικό αριθμό των υπαλλήλων είναι ταξινομημένο κατά τον κωδικό αριθμό των υπαλλήλων και ότι ζητάμε να εξάγουμε από αυτό τις λογικές εγγραφές όλων των υπαλλήλων που ασχολούνται στο τμήμα της επιχείρησης με κωδικό **8**. Και στην περίπτωση αυτή πρέπει να ερευνηθεί ολόκληρο το αρχείο είτε βρίσκεται στην μαγνητοταινία είτε στον δίσκο.

γ) Εξαγωγική ανάληψη με κλειδί το ίδιο με αυτό με το οποίο έχει ταξινομηθεί το αρχείο. Για παράδειγμα, εξαγωγή της λογικής εγγραφής του υπαλλήλου με κωδικό αριθμό **23456**. Η διαδικασία εξαγωγικής ανάληψης του τύπου αυτού απαιτεί έρευνα των λογικών εγγραφών του αρχείου μέχρις ότου ευρεθεί η ζητούμενη.

Επάνω σε ένα σειριακό αρχείο μπορεί να εφαρμοστεί η "σειριακή έρευνα" (**SEQUENTIAL SEARCH**), δηλαδή να ερευνηθεί κανείς μια προς μια τις λογικές εγγραφές κατά σειρά αρχής γενομένης από την πρώτη. Στην περίπτωση αυτή η διαδικασία της αναζήτησης μπορεί να σταματήσει στην πρώτη λογική εγγραφή (αν η ζητούμενη είναι η πρώτη του αρχείου) είτε στην τελευταία (αν η ζητούμενη είναι η τελευταία). Εάν το αρχείο είναι ταξινομημένο, τότε μπορεί να εφαρμοστεί η "δυναμική αναζήτηση" (**BINARY SEARCH**), η οποία εν γένει απαιτεί έλεγχο λιγότερων λογικών εγγραφών του αρχείου. Η μέθοδος αυτή, η οποία αξιοποιεί την δυνατότητα που υπάρχει στον δίσκο για απ'ευθείας προσπέλαση σε μια φυσική εγγραφή, διχοτομεί διαδοχικά το αρχείο και εντοπίζει το τμήμα στο οποίο ανήκει η ζητούμενη λογική εγγραφή. Δεν ανήκει στις προθέσεις μας η λεπτομερής ανάπτυξη της μεθόδου, ωστόσο είναι ενδιαφέρον να σημειωθεί ότι ένα σειριακό αρχείο στον δίσκο με **2000** εγγραφές απαιτεί με την μέθοδο αυτή το πολύ **11** εγγραφές του να εισαχθούν στην μνήμη και να ελεγχθούν. Πιο κάτω φαίνεται ένα παράδειγμα της μεθόδου με αρχείο που περιλαμβάνει **9** λογικές εγγραφές. Για την

απλοποίηση της παρουσίασης κάθε φυσική εγγραφή περιλαμβάνει μόνο μια λογική εγγραφή και έστω ότι αναζητούμε την εγγραφή με κωδικό **500**.

αρχείο:	235	400	403	450	460	500	507	600	624
---------	------------	------------	------------	------------	------------	------------	------------	------------	------------

Εισάγουμε την μεσαία εγγραφή του αρχείου στην μνήμη και συγκρίνουμε τον κωδικό της με τον ζητούμενο κωδικό. Επειδή ο κωδικός της μεσαίας εγγραφής (**460**) είναι μικρότερος του ζητούμενου (**500**) έπεται ότι ο δεύτερος βρίσκεται στο τμήμα του αρχείου μεταξύ **460** και **624**. Εισάγουμε τώρα την μεσαία εγγραφή του τμήματος αυτού (**507**). Επειδή ο κωδικός της είναι μεγαλύτερος του **500** έπεται ότι ο ζητούμενος κωδικός βρίσκεται στο τμήμα μεταξύ **460** και **507**. Εισάγουμε τον μεσαίο του τμήματος αυτού (**500**) και συγκρίνουμε τον κωδικό του με τον ζητούμενο. Επειδή υπάρχει σύμπτωση έπεται ότι η εγγραφή που εισήχθητε τελευταία στην μνήμη είναι η ζητούμενη.

3.1.4 ΟΡΓΑΝΩΣΗ ΑΡΧΕΙΩΝ ΑΜΕΣΗΣ ΠΡΟΣΠΕΛΑΣΗΣ

Άμεση προσπέλαση σε ένα αρχείο σημαίνει την δυνατότητα ανάληψης μιας λογικής εγγραφής του απευθείας χωρίς να απαιτείται έρευνα των υπολοίπων λογικών εγγραφών του αρχείου. Βέβαια την δυνατότητα αυτή παρέχουν μόνο τα μέσα βοηθητικής μνήμης που από την φύση τους επιτρέπουν την άμεση προσπέλαση σε οποιοδήποτε τμήμα τους. Έτσι, τα αρχεία άμεσης προσπέλασης καταχωρούνται σε μέσα άμεσης προσπέλασης. Το μέσο που έχει επικρατήσει να χρησιμοποιείται ευρέως είναι ο μαγνητικός δίσκος.

Κατά κανόνα, το **SOFTWARE** συστήματος των υπολογιστών που διαθέτουν δίσκους είναι εφοδιασμένο με ειδικό "σύστημα χειρισμού αρχείων" (**FILE MANAGEMENT SYSTEM-FMS**). Το σύστημα αυτό επιτρέπει στο χρήστη να δημιουργήσει ένα αρχείο στον δίσκο και έπειτα να ανασύρει από αυτό μια

οποιαδήποτε λογική εγγραφή. Το σύστημα αριθμεί τις λογικές εγγραφές του αρχείου από 1 έως N-αν N είναι το πλήθος των λογικών εγγραφών του αρχείου-και εντοπίζει την διεύθυνση μιας λογικής εγγραφής στον δίσκο(επιφάνεια, άτρακτος, τομέας) από τον αύξοντα αριθμό της. Έτσι ο χρήστης δεν ασχολείται με την επιφάνεια , την άτρακτο και τον τομέα που βρίσκεται μια λογική εγγραφή. Από την σκοπιά του χρήστη η φυσική διεύθυνση μιας λογικής εγγραφής είναι ο αύξων αριθμός της. Και γεννάται τώρα το ερώτημα: Πως θα πρέπει να οργανώσει ο χρήστης το αρχείο του έτσι ώστε να συσχετίσει τις διευθύνσεις των λογικών εγγραφών με το περιεχόμενό τους; Πως θα εντοπίσει την λογική εγγραφή ενός πελάτη με κωδικό **23574** ή την λογική εγγραφή του αποθέματος με κωδικό **455**;

Ένας τρόπος είναι να ερευνήσει το αρχείο εγγραφή προς εγγραφή έως ότου εντοπίσει το ζητούμενο, αλλά αυτό είναι κάτι που γίνεται στην περίπτωση της σειριακής οργάνωσης και δεν έχει το πλεονέκτημα της άμεσης προσπέλασης στο ζητούμενο.

3.1.4.1. ΜΕΘΟΔΟΙ ΣΥΣΧΕΤΙΣΜΟΥ ΚΛΕΙΔΙΟΥ ΚΑΙ ΔΙΕΥΘΥΝΣΗΣ

Πιο κάτω αναφέρονται σε συντομία μερικές από τις μεθόδους συσχετισμού κλειδιού και διεύθυνσης των λογικών εγγραφών.

α)Κλειδί και διεύθυνση ίδια

Με την μέθοδο αυτή το κλειδί των λογικών εγγραφών συμπίπτει με την διεύθυνσή τους. Για παράδειγμα, αποθηκεύουμε την λογική εγγραφή του υλικού αποθήκης με κωδικό **9** στην διεύθυνση **9**. Αλλά αυτό είναι τις περισσότερες φορές αδύνατο, επειδή, οι κωδικοί των υλικών σπάνια συμπίπτουν με τον αύξοντα αριθμό τους.

β) Πίνακας κλειδιού και διεύθυνσης

Η μέθοδος αυτή είναι η πιο συνηθισμένη και συνίσταται στο να συνταχθεί ένας πίνακας που να αντιστοιχεί το κλειδί με την διεύθυνση στην οποία έχει καταχωρηθεί η λογική εγγραφή. Στο σχήμα 4.2 φαίνεται ένας τέτοιος πίνακας για ένα αρχείο πελατών.

ΚΛΕΙΔΙ (κωδικός πελάτη)	ΔΙΕΥΘΥΝΣΗ (αριθμός εγγραφής)
44356	530
44550	22
44600	1310
.	.
.	.
.	.
56345	299
56450	300

Σχήμα3-2: Πίνακας Κλειδιών-Διευθύνσεων

Για την ανεύρεση του πελάτη με κωδικό **44600** διερευνάται ο πίνακας και εντοπίζεται η διεύθυνση με την οποία το σύστημα χειρισμού αρχείων εισάγει την αντίστοιχη λογική εγγραφή στην μνήμη για τα περαιτέρω.

Ο πίνακας αυτός συνήθως φυλάσσεται στον δίσκο, από τον οποίο εισάγεται στην μνήμη για αναζήτηση. Η έρευνα του πίνακα γίνεται είτε εξαντλητικά ελέγχοντας όλους τους κωδικούς έως ότου εντοπιστεί ο ζητούμενος είτε, περισσότερο αποτελεσματικά με δυαδική αναζήτηση.

Γ) Υπολογισμός διεύθυνσης από το κλειδί

Στη μέθοδο αυτή αντιστοιχεί μια διεύθυνση σε ένα κλειδί με κάποιο υπολογισμό. Για παράδειγμα, αν το κλειδί είναι το επώνυμο των πελατών, τότε στο επώνυμο

ΠΑΠΠΙΑΣ μπορεί να αντιστοιχηθεί η διεύθυνση που προκύπτει ως άθροισμα των αριθμών σειράς στην αλφάβητο των γραμμάτων που απαρτίζουν το επώνυμο αυτό:

$$16+1+16+16+1+18 = 68$$

Ένας άλλος υπολογισμός που εφαρμόζεται πολύ συχνά για αριθμητικά κλειδιά είναι να αντιστοιχισθεί η διεύθυνση που προκύπτει ως υπόλοιπο της διαίρεσης του κλειδιού με τον μεγαλύτερο πρώτο αριθμό που δεν ξεπερνά το μέγεθος του αρχείου. Για παράδειγμα, προκειμένου περί αρχείου με **1000** λογικές εγγραφές ο διαιρέτης θα είναι **997** και επομένως στο κλειδί **3722** θα αντιστοιχισθεί η διεύθυνση **731(3722 = 997 * 3 + 731)**

Η υπολογιστική μέθοδος έχει το μειονέκτημα ότι είναι δυνατό να παραχθεί η ίδια διεύθυνση για δύο διαφορετικά κλειδιά και έτσι να έχουμε δύο "συνώνυμες" λογικές εγγραφές. Για παράδειγμα στον προηγούμενο υπολογισμό τόσο το κλειδί **3722** όσο και το κλειδί **2725** παράγουν την ίδια διεύθυνση **731**.

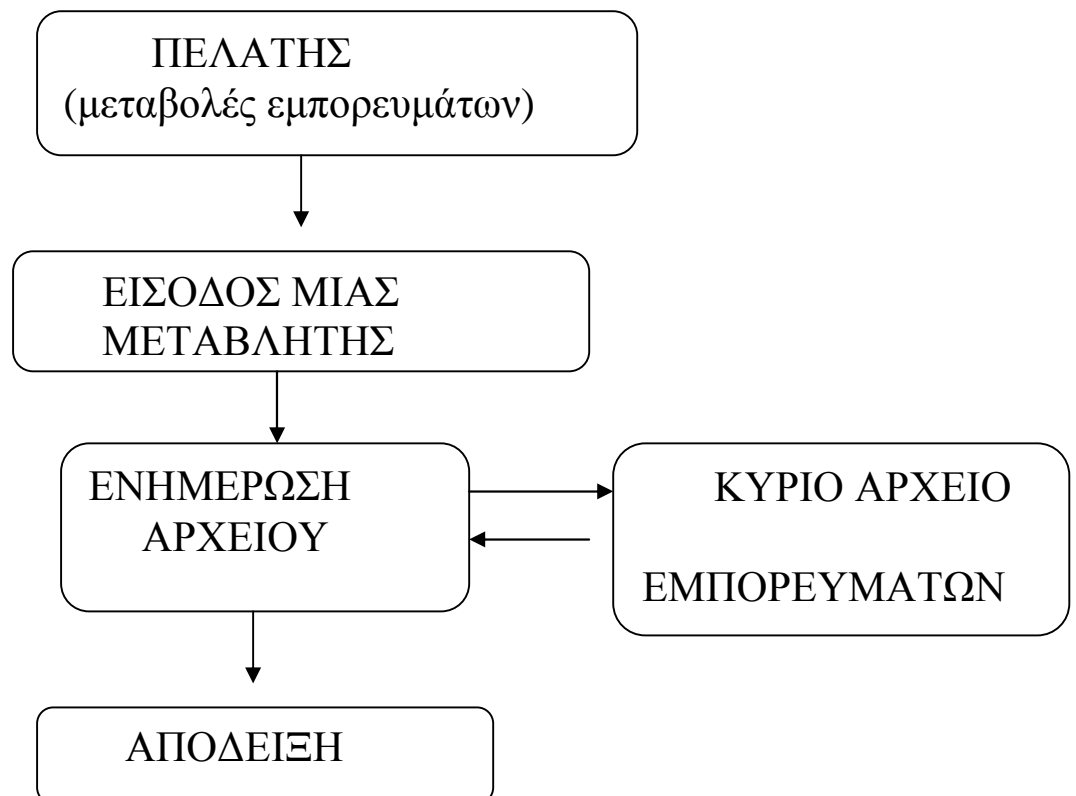
3.1.4.2. ΕΝΗΜΕΡΩΣΗ ΑΡΧΕΙΟΥ ΑΜΕΣΗΣ ΠΡΟΣΠΕΛΑΣΗΣ

Ένα αρχείο άμεσης προσπέλασης μπορεί να ενημερωθεί όπως και ένα σειριακό αρχείο κατά τρόπο ομαδικό (**BATCH PROCESSING**) αλλά το μεγάλο πλεονέκτημά του είναι ότι μπορεί να ενημερωθεί **ON-LINE** (άμεσα) μόλις συμβεί μεταβολή έστω και σε μια λογική εγγραφή του.

Έστω για παράδειγμα το αρχείο των εμπορευμάτων ενός **Super Market** με τον πίνακα κωδικών και διευθύνσεων των εμπορευμάτων. Το αρχείο βρίσκεται **ON-LINE** σε δίσκο το δε ταμείο είναι μια τερματική οθόνη με πληκτρολόγιο. Το όλο σύστημα διαθέτει ένα εκτυπωτή για την έκδοση τιμολογίου. Κάθε φορά που ένας πελάτης φθάνει στο ταμείο η ταμίας εισάγει από το τερματικό τις ποσότητες των εμπορευμάτων που αγόρασε ο πελάτης και το αρχείο των εμπορευμάτων

ενημερώνεται, μέσω κατάλληλου προγράμματος. Συγχρόνως εκδίδεται και απόδειξη από τον εκτυπωτή.

Στο Σχήμα3.3 φαίνεται διαγραμματικά η ενημέρωση αυτή.



Σχήμα 3-3:Ενημέρωση αρχείου άμεσης προσπέλασης

Κατά την ενημέρωση αυτή, εντοπίζεται η λογική εγγραφή του εμπορεύματος στο αρχείο από τον κωδικό του, με την βοήθεια του πίνακα κωδικών-διευθύνσεων, εισάγεται στην μνήμη, υφίσταται ενημέρωση με την αφαίρεση της ποσότητας του εμπορεύματος που αγοράστηκε από το απόθεμα και η ενημέρωση πλέον λογική εγγραφή ξανά καταχωρείται στην διεύθυνσή της, σβήνοντας την προηγούμενη.

3.1.4.3 ΑΝΑΛΗΨΗ ΠΛΗΡΟΦΟΡΙΩΝ ΑΠΟ ΤΟ ΑΡΧΕΙΟ

α) **Επιλεκτική πληροφόρηση**, όπως η εκτύπωση πίνακα υλικών των οποίων το απόθεμα είναι κάτω του ορίου ασφαλείας. Στο παράδειγμα αυτό θα απαιτηθεί έλεγχος όλων των λογικών εγγραφών του κύριου αρχείου αποθεμάτων και σύγκριση για κάθε μια του τρέχοντος αποθέματος με το απόθεμα ασφαλείας κάθε εμπορεύματος. Γίνεται φανερό ότι η πληροφόρηση αυτή θα γίνει πολύ ταχύτερα με σειριακή επεξεργασία του αρχείου από την αρχή μέχρι το τέλος.

β) **Εξαγωγική πληροφόρηση**, όπως η εκτύπωση όλων των στοιχείων των υπαλλήλων ενός οργανισμού με μισθολογικό κλιμάκιο **8**. Στην περίπτωση αυτή θα μπορούσε να έχει υπάρξει κατά την δημιουργία του αρχείου ένας πίνακας κλιμακίων και διευθύνσεων. Έτσι, εισάγεται ο πίνακας στην μνήμη και εξετάζεται. Κάθε φορά που θα συναντάται το κλιμάκιο **8** θα εισάγεται από το αρχείο η λογική εγγραφή της αντίστοιχης διεύθυνσης και θα εκτυπώνεται.

3.2 ΣΥΝΘΕΤΟΙ ΜΕΘΟΔΟΙ ΟΡΓΑΝΩΣΗΣ ΑΡΧΕΙΩΝ

Πέρα από τις δύο βασικές αναφερθείσες οργανώσεις αρχείων, την σειριακή και την άμεσης προσπέλασης, έχουν αναπτυχθεί και άλλοι τρόποι που συνδυάζουν ορισμένα πλεονεκτήματα και των δύο και επιτρέπουν ταχεία επεξεργασία σε ειδικές περιπτώσεις εφαρμογών. Θα αναφερθούμε σύντομα στους τρόπους αυτούς.

Ας σημειωθεί ότι στο παράδειγμα αυτό έχουμε μόνο ένα κλειδί και ένα δείκτη. Θα μπορούσαμε να είχαμε δύο κλειδιά και δύο δείκτες (π.χ όνομα και ηλικία) και επομένως δύο διαφορετικές ακολουθίες.

Να σημειωθεί ακόμα ότι η οργάνωση αυτή παρουσιάζει προβλήματα κατά την διαγραφή εγγραφών ή προσθήκη νέων στο αρχείο και τούτο επειδή πρέπει να ενημερωθούν οι δείκτες.

β)Ανεστραμμένος Κατάλογος(INVERTED LIST)

Όπως είδαμε στον απλό κατάλογο προηγουμένως ,μπορεί να έχουμε περισσότερες της μιας ακολουθίες ,κάθε μια από τις οποίες αντιστοιχεί σε ένα κλειδί.

Θα μπορούσε λοιπόν να συνταχθεί ένας πίνακας που στην μια στήλη του να περιλαμβάνει τα διαφορετικά αυτά κλειδιά και στην άλλη τις τιμές των δεικτών των πρώτων λογικών εγγραφών των ακολουθιών, όπως φαίνεται στο σχήμα 4.5.

ΜΗΧΑΝΙΚΟΣ	50	ΕΝΑΡΚΤΗΡΙΕΣ
ΓΙΑΤΡΟΣ	48	ΤΙΜΕΣ
ΦΥΣΙΚΟΣ	103	ΔΕΙΚΤΩΝ

Σχήμα 3-5: Πίνακας Κλειδιών –Εναρκτήριων Τιμών Δεικτών

Με την χρησιμοποίηση του πίνακα αυτού γίνεται φανερό ότι τα κλειδιά του πίνακα είναι πληροφορίες που θα μπορούσαν να απομακρυνθούν από τις λογικές εγγραφές αφού η ύπαρξη της λογικής εγγραφής στην ακολουθία

συνεπάγεται το χαρακτηριστικό αυτό. Μια περίπτωση είναι ένα αρχείο υπαλλήλων. Η οργάνωση αυτή καλείται **“μερικά ανεστραμμένος κατάλογος” (PARTIALLY INVERTED LIST)**.

Η ιδέα αυτή μπορεί να επεκταθεί περιλαμβάνοντας όλα τα πεδία των λογικών εγγραφών στον πίνακα και τοποθετώντας τις διευθύνσεις των λογικών εγγραφών, στις οποίες περιλαμβάνονται. Έτσι, έχουμε ένα πλήρως ανεστραμμένο αρχείο, του οποίου οι λογικές εγγραφές δεν χρειάζεται να περιέχουν τα χαρακτηριστικά που υπάρχουν στον πίνακα αλλά ούτε και δείκτες. Στο παρακάτω σχήμα φαίνεται ένα σχετικό παράδειγμα.

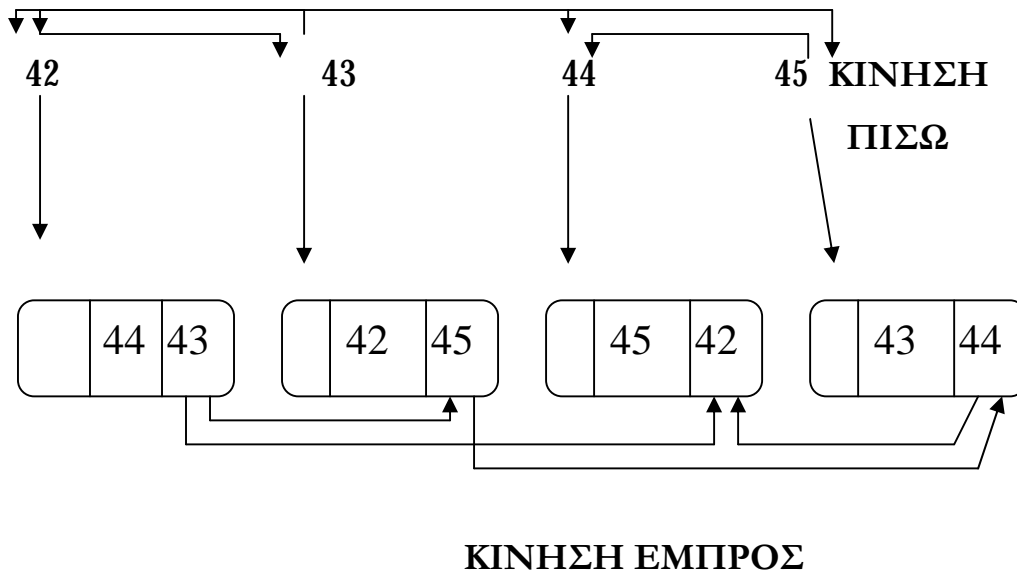
ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ	ΔΙΕΥΘΥΝΣΕΙΣ ΕΓΓΡΑΦΩΝ	<u>ΜΟΝΟ ΟΝΟΜΑΤΑ</u>
ΜΗΧΑΝΙΚΟΣ	42,43,45	42 ΠΑΠΠΑΣ Γ.
ΓΙΑΤΡΟΣ	44,46	43 ΑΝΤΩΝΙΟΥ Χ
ΑΘΗΝΑ	42,44	44 ΒΑΣΟΣ Δ.
ΛΑΜΙΑ	43	45 ΧΡΗΣΤΟΥ Χ.
ΠΑΤΡΑ	45	46 ΓΕΩΡΓΙΟΥ
ΠΥΡΓΟΣ	46	
.	.	
.	.	
.	.	

Σχήμα 3-7: Πλήρως Ανεστραμμένος Κατάλογος

γ) Δακτύλιος

Η οργάνωση δακτυλίου περιλαμβάνει σε κάθε λογική εγγραφή του αρχείου δύο δείκτες καταλόγου, ένα που να δείχνει την επόμενη εγγραφή και ένα που να δείχνει την προηγούμενη. Έτσι, είναι δυνατή η κίνηση διαμέσου των

λογικών εγγραφών του αρχείου τόσο προς τα εμπρός όσο και προς τα πίσω. Η τελευταία εγγραφή του αρχείου έχει ένδειξη της πρώτης. Στο σχήμα 4.8 φαίνεται διαγραμματικά η οργάνωση αυτή.



Σχήμα 3-8: Δακτύλιος (πρώτος δείκτης = προηγούμενη εγγραφή, δεύτερος δείκτης = επόμενη εγγραφή)

3.2.2 ΠΙΝΑΚΟΠΟΙΗΜΕΝΗ ΣΕΙΡΙΑΚΗ ΟΡΓΑΝΩΣΗ (INDEXED SEQUENTIAL ACCESS METHOD)

Η Πινακοποιημένη Σειριακή Οργάνωση (INDEXED SEQUENTIAL ACCESS METHOD-ISAM) συνδυάζει τα πλεονεκτήματα, ως προς την ενημέρωση ενός αρχείου, της σειριακής οργάνωσης και τα πλεονεκτήματα ως προς την ανάληψη πληροφοριών της οργάνωσης άμεσης προσπέλασης. Κατά την οργάνωση αυτή οι λογικές εγγραφές καταχωρούνται στο δίσκο σειριακά. Επί πλέον, δημιουργείται πίνακας κλειδιών-διευθύνσεων.

Η ανάληψη μιας λογικής εγγραφής γίνεται άμεσα με εξέταση του πίνακα.

Κατά την ενημέρωση του αρχείου οι λογικές εγγραφές που εμφανίζουν μεταβολές εντοπίζονται από τον πίνακα, εισάγονται στην μνήμη, υφίστανται

την ενημέρωση και επανεγγράφονται στην ίδια διεύθυνσή τους. Η προσθήκη μιας νέας λογικής εγγραφής δεν γίνεται στην σειρά της αλλά στο τέλος ,ενώ τοποθετείται ένας δείκτης στην λογικά προηγούμενη εγγραφή που προσδιορίζει την διεύθυνσή της. Η διαγραφή μιας λογικής εγγραφής αφήνει κενή θέση. Το αρχείο περιοδικά ξαναοργανώνεται έτσι ώστε να τοποθετούνται όλες οι λογικές εγγραφές στην σειρά τους και να απαλείφονται τα κενά. Η οργάνωση αυτή χρησιμοποιήθηκε και χρησιμοποιείται ευρύτατα. Τα περισσότερα συστήματα υπολογιστών περιλαμβάνουν στο **SOFTWARE** συστήματος ειδικές ρουτίνες που υποβοηθούν τον χρήστη στη δημιουργία και επεξεργασία τέτοιων αρχείων. Στα προγράμματα εφαρμογής ο χρήστης απλά καθορίζει ένα κλειδί και οι ρουτίνες συστήματος αναλαμβάνουν να εντοπίσουν την ζητούμενη λογική εγγραφή.

3.3. ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ

Τα απλά Συστήματα Επεξεργασίας Αρχείων ικανοποιούν πάρα πολλές απαιτήσεις σε επιχειρήσεις και οργανισμούς και κυρίως σε περιπτώσεις όπου υπάρχει μικρός αριθμός χρηστών και τα δεδομένα μπορούν να κατανεμηθούν εύκολα χωρίς να απαιτείται δημιουργία Ολοκληρωμένων Πληροφοριακών Συστημάτων. Σε άλλες όμως περιπτώσεις , όπως σε ορισμένα συστήματα συναλλαγών, στα Πληροφοριακά Συστήματα Διοίκησης και στα Συστήματα Υποστήριξης Αποφάσεων, τα συστήματα επεξεργασίας αρχείων δεν μπορούν να ικανοποιήσουν τις απαιτήσεις , καθώς όλα αυτά τα συστήματα περιλαμβάνουν μεγάλο αριθμό χρηστών, πολλά αρχεία με δεδομένα που σχετίζονται μεταξύ τους και διαφορετικές όψεις των ίδιων δεδομένων. Ακόμη περισσότερο, χρήστες αυτών των συστημάτων απαιτούν γρήγορες απαντήσεις σε μη προσχεδιασμένες ερωτήσεις, μια λειτουργία που δεν υποστηρίζεται από τα απλά συστήματα επεξεργασίας αρχείων.

Τα στοιχεία που χαρακτηρίζουν μια **Βάση Δεδομένων** είναι:

1. Περιορισμός της πολλαπλής αποθήκευσης στοιχείων(**Redundancy**)
2. Μοίρασμα των πληροφοριών σε πολλούς χρήστες(**Sharing**)
3. Ενιαίος τρόπος χειρισμού των πληροφοριών(**Uniformity**)
4. Ασφάλεια των πληροφοριών(**Security**)
5. Ακεραιότητα των πληροφοριών(**Integrity**)
6. Ανεξαρτησία των δεδομένων και των εφαρμογών από τον τρόπο αποθήκευσης (**Data Independence**)
7. Ευκολία στη χρήση.

3.3.1 ΣΥΣΤΗΜΑΤΑ ΔΙΑΧΕΙΡΙΣΗΣ ΒΑΣΕΩΝ ΔΕΔΟΜΕΝΩΝ

Τα Συστήματα Διαχείρισης Βάσεων Δεδομένων(ΣΔΒΔ), αγγλικός όρος **Data Base Management Systems(DBMS)**, είναι προγράμματα τα οποία ακολουθούν μια πιο πολύπλοκη προσέγγιση στην αποθήκευση και ανάκτηση δεδομένων από αυτή που ακολουθούν τα απλά Συστήματα Επεξεργασίας Αρχείων(**File Management System-FMS**). Στην πραγματικότητα τα ΣΔΒΔ αρχίζουν εκεί που σταματούν τα απλά συστήματα επεξεργασίας αρχείων. Τα ΣΔΒΔ εκμεταλλεύονται την ταχύτητα και την ευελιξία που προσφέρει η μέθοδος της άμεσης προσπέλασης ,όσον αφορά την αποθήκευση και την ανάκτηση δεδομένων, οργανώνουν τα δεδομένα με πολύ διαφορετικό τρόπο από ότι τα απλά συστήματα. Ένα ΣΔΒΔ είναι απλά το λογισμικό το οποίο υποστηρίζει την συγκέντρωση των δεδομένων, την αποτελεσματική διαχείρισή τους και επιτρέπει την πρόσβαση των εφαρμογών στις αποθηκευμένες αυτές πληροφορίες.

Ένα από τα μεγαλύτερα πλεονεκτήματα των ΣΔΒΔ είναι ότι το σύστημα βρῖσκει και ανακτά τα δεδομένα χωρίς τη μεσολάβηση του χρήστη. Ακόμη και οι προγραμματιστές Η/Υ δεν απαιτείται να γνωρίζουν την τοποθεσία

και την μορφή των εγγραφών. Απλά ζητούν να ανευρεθούν τα δεδομένα μέσω των προγραμμάτων και το σύστημα αναλαμβάνει να τα ανακτήσει για λογαριασμό τους. Τα συστήματα επεξεργασίας συμβατικών αρχείων όμως λειτουργούν διαφορετικά και με τρόπο που μοιάζει με την αναζήτηση ενός βιβλίου σε μια βιβλιοθήκη, όπου βρίσκουμε κάποιο βιβλίο είτε ψάχνοντας τα ράφια (επεξεργασία μέσω σειριακής αναζήτησης), είτε ψάχνοντας στις καρτέλες της βιβλιοθήκης (επεξεργασία μέσω ευρετηρίου και άμεσης προσπέλασης). Όποιον τρόπο και από τους δύο και αν επιλέξουμε, θα πρέπει μόνοι μας να πάμε στα ράφια της βιβλιοθήκης για να πάρουμε το βιβλίο. Αντίστοιχα, με ένα ΣΔΒΔ είναι σαν να ζητούμε από τον υπάλληλο της βιβλιοθήκης να βρεί το βιβλίο που επιθυμούμε και να μας το φέρει προσδιορίζοντας το τίτλο, το συγγραφέα και το έτος έκδοσης. Η διαδικασία που ακολουθεί ο υπάλληλος και ο τρόπος οργάνωσης των βιβλίων στα ράφια δεν μας ενδιαφέρουν, δηλαδή δεν μας ενδιαφέρει αν ο υπάλληλος χρησιμοποίησε τις καρτέλες της βιβλιοθήκης για να βρει τη θέση του βιβλίου ή αν τα βιβλία είναι οργανωμένα στα ράφια βάσει του τίτλου, του συγγραφέα, της ημερομηνίας έκδοσης ή του μεγέθους τους. Ακόμη περισσότερο μπορούμε με την ίδια ευκολία να ζητήσουμε από τον υπάλληλο να μας φέρει όλα τα βιβλία ενός συγκεκριμένου συγγραφέα ή τα άρθρα που δημοσιεύτηκαν πάνω σε ένα συγκεκριμένο θέμα την τελευταία πενταετία χωρίς να μας ενδιαφέρει πως θα γίνει αυτό.

Αν και τα ΣΔΒΔ παρέχουν ταχύτητα και αξιοπιστία μεγαλύτερη από τα συμβατικά συστήματα, συχνά δεν χρησιμοποιούνται εξαιτίας του κόστους που τα συνοδεύει. Ένας άλλος λόγος είναι ότι τα συστήματα αναπτύχθηκαν πριν την εμφάνιση της τεχνολογίας των Βάσεων Δεδομένων και οι επιχειρήσεις ή οργανισμοί δείχνουν απροθυμία να αλλάξουν τον τρόπο λειτουργίας τους. Ένας ακόμη λόγος είναι η πολυπλοκότητα ανάπτυξης και συντήρησης ενός τέτοιου συστήματος. Η απόφαση ανάπτυξης ενός ΣΔΒΔ πρέπει να είναι απόρροια μιας ανάλυσης κόστους / οφέλους, καθώς τόσο τα

ΣΔΒΔ όσο και τα απλά συστήματα επεξεργασίας αρχείων είναι κατάλληλα για διαφορετικές περιπτώσεις. **Ο γενικός κανόνας που πρέπει να ακολουθείται στην επιλογή συστήματος είναι να ικανοποιούνται οι απαιτήσεις των χρηστών με το μικρότερο δυνατό κόστος.**

Υπάρχουν τέσσερα συστατικά στον ορισμό της Βάσης Δεδομένων τα οποία πρέπει να τονιστούν ιδιαίτερα. Πρώτον, υπάρχει **η φυσική υπόσταση των δεδομένων**, τα οποία τηρούνται σε κάποιο μαγνητικό μέσο, συνήθως σε μαγνητικούς δίσκους. Στη Βάση Δεδομένων υπάρχουν τα δεδομένα που χρησιμοποιούνται από κάθε λειτουργική μονάδα του οργανισμού ή της επιχείρησης όπως π.χ. το Προσωπικό. Εν αντιθέσει με τα συμβατικά συστήματα, τα δεδομένα στα ΣΔΒΔ αποθηκεύονται σε μια τοποθεσία, προσδιορίζονται οριστικά και αμετάκλητα και χρησιμοποιούνται από όλες τις εφαρμογές που έχουν πρόσβαση στην συγκεκριμένη τοποθεσία.

Το δεύτερο συστατικό ενός Συστήματος Διαχείρισης Βάσεις Δεδομένων είναι το Λογισμικό που χρησιμοποιείται διαμεσολαβητικά ανάμεσα στις εφαρμογές και στα φυσικά δεδομένα και τα αρχεία όπου αυτά είναι αποθηκευμένα. Όταν μια εφαρμογή ζητήσει κάποια πληροφορία, το σύστημα Διαχείρισης Βάσεων Δεδομένων την βρίσκει και την μεταφέρει στην εφαρμογή χωρίς ο προγραμματιστής ή ο χρήστης να απαιτείται να προσδιορίσει την συγκεκριμένη πληροφορία ή να πει στο σύστημα που βρίσκεται. Όλες οι εντολές αναζήτησης που χρησιμοποιούνται στις συμβατικές γλώσσες προγραμματισμού εδώ είναι περιττές. Τα συστήματα Διαχείρισης Βάσεων Δεδομένων προσφέρουν στους προγραμματιστές μια γλώσσα προσδιορισμού δεδομένων η οποία μπορεί να χρησιμοποιηθεί για να παρέχει την σύνδεση μεταξύ εφαρμογών και αρχείων.

Το τρίτο συστατικό είναι ένα σύνολο εφαρμογών όπου χρησιμοποιούνται για την ανεύρεση δεδομένων στη Βάση. Οι εφαρμογές αυτές είναι συνήθως γραμμένες σε κάποια συμβατική γλώσσα όπως η **COBOL** και τα δεδομένα, τα οποία απαιτούνται μέσω μιας εφαρμογής σε **COBOL**,

ανευρίσκονται και μεταφέρονται στο σύστημα Διαχείρισης Βάσεων Δεδομένων χωρίς ο χρήστης να έχει προσδιορίσει ούτε πως ούτε πού μπορεί να βρῖσκειται η πληροφορία αυτή.

Ένα Σύστημα Διαχείρισης Βάσης Δεδομένων συνοδεύεται από ένα πακέτο λογισμικού το οποίο σε γενικές γραμμές περιλαμβάνει μια γλώσσα προγραμματισμού δεδομένων και μια γλώσσα διαχείρισης αυτών. Η γλώσσα προσδιορισμού των δεδομένων είναι η γλώσσα που χρησιμοποιείται από τους προγραμματιστές για την ανάπτυξη της Βάσης Δεδομένων.

Τα περισσότερα Συστήματα Διαχείρισης Βάσης Δεδομένων έχουν την δική τους γλώσσα για την ανάπτυξη εφαρμογών και η οποία αποκαλείται γλώσσα διαχείρισης δεδομένων. Αυτή η γλώσσα είναι τέταρτης γενιάς και φιλική προς τον χρήστη παρέχοντας έναν αριθμό εντολών που προσφέρουν τόσο στους χρήστες όσο και στους προγραμματιστές τις απαιτούμενες δυνατότητες για την διαχείριση των δεδομένων. Η πλέον σημαντική από τις γλώσσες αυτές είναι η **Γλώσσα Δομημένων Ερωτήσεων (Structured Query Language-SQL)**.

Η χρησιμοποίηση της γλώσσας Διαχείρισης Δεδομένων ενός Συστήματος Διαχείρισης Βάσης Δεδομένων επιτρέπει την συγγραφή προγραμμάτων από τους τελικούς χρήστες (υπάλληλοι, διευθυντές χωρίς εξειδικευμένη γνώση). Τα προγράμματα αυτά περιέχουν μερικές απλές ερωτήσεις με σκοπό την παραγωγή κάποιας αναφοράς.

Το τέταρτο συστατικό ενός Συστήματος Διαχείρισης Βάσης Δεδομένων είναι το λεξικό Δεδομένων το οποίο δεν είναι παρά ένα αυτοματοποιημένο αρχείο όπου τηρούνται οι ορισμοί των δεδομένων και τα χαρακτηριστικά τους όπως χρήση, φυσική αναπαράσταση ,ιδιοκτησία(ποιος είναι υπεύθυνος για την συντήρησή τους), δικαιώματα πρόσβασης και ασφάλεια. Σε πολλές περιπτώσεις αυτά τα Λεξικά μπορούν να παράγουν αναφορές που

καταγράφουν αναλυτικά την χρήση των στοιχείων, τις τυχόν ομαδοποιήσεις των δεδομένων, εφαρμογές που έχουν πρόσβαση σε αυτά κ.λπ.

Τα περισσότερα όμως Λεξικά είναι εντελώς παθητικά ,δεν παρέχουν δηλαδή πληροφορίες όπως αυτές που προαναφέρθηκαν. Υπάρχουν όμως τα Λεξικά τα οποία είναι αρχικά προηγμένα και τα οποία φροντίζουν μετά από κάποια μεταβολή στο περιεχόμενό τους να ενημερώνουν όλες τις εφαρμογές που σχετίζονται με αυτή τη μεταβολή. Για παράδειγμα, αν χρειαστεί να αλλάξουμε τους Ταχυδρομικούς Κώδικες από πέντε σε εννιά ψηφία, απλώς εισάγουμε στο Λεξικό την μεταβολή αυτή και μετά αυτό αναλαμβάνει να ενημερώσει όλες τις εφαρμογές οι οποίες κάνουν χρήση Ταχυδρομικών Κωδικών.

3.3.2 ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΣΥΣΤΗΜΑΤΩΝ ΒΑΣΕΩΝ ΔΕΔΟΜΕΝΩΝ

Η **αρχιτεκτονική** η οποία πολλές φορές αναφέρεται με την ονομασία **ANSI SPARC**, όπου παρατηρούμε ότι υπάρχουν πολλοί χρήστες , που μπορεί να είναι προγραμματιστές μέχρι και απλοί χρήστες, που προσπελάνουν την Βάση Δεδομένων με προγράμματα είτε με **επεξεργασία (batch)** είτε **ανοικτής γραμμής (on-line)**.

Κάθε χρήστης έχει στην διάθεσή του μια γλώσσα η οποία μπορεί να είναι είτε **μια γλώσσα ερωτήσεων (query language)** είτε μια συμβατική γλώσσα (π.χ. **COBOL** ή **FORTRAN**) η οποία όμως έχει επεταθεί κατά τέτοιο τρόπο ώστε να μπορεί να διαχειρίζεται την Βάση Δεδομένων. Η **συμβατική γλώσσα** ονομάζεται **φιλόξενη (host)** ενώ η επέκτασή της που της επιτρέπει την διαχείριση λέγεται **Υπογλώσσα Δεδομένων (ΥΔ)**, **αγγλικός όρος Data sub-Language (DSL)**.

Η **Υπογλώσσα Δεδομένων** αποτελείται από δύο τμήματα: τη **γλώσσα ορισμού δεδομένων** και τη **γλώσσα διαχείρισης δεδομένων** στις οποίες έχουμε ήδη αναφερθεί. Τις περισσότερες φορές η γλώσσα ορισμού δεδομένων χρησιμοποιείται ,εξωτερικά, από το πρόγραμμα ενώ η γλώσσα διαχείρισης δεδομένων είναι μια σειρά από κλήσεις του ΣΔΒΔ. Κάθε χρήστης έχει στην διάθεσή του μια περιοχή εργασίας (**work space**) όπου τηρούνται όλες οι πληροφορίες της επικοινωνίας με τη Βάση Δεδομένων.

Παρατηρούμε λοιπόν ότι στην αρχιτεκτονική ο κάθε χρήστης ή ομάδα χρηστών έχουν μια **εξωτερική άποψη (external view)** της Βάσης Δεδομένων. Κάθε διακεκριμένη εξωτερική όψη αποκαλείται **εξωτερικό σχήμα(external schema)** και περιγράφεται μέσω της γλώσσας διαχείρισης δεδομένων η οποία και διαχειρίζεται τις **εξωτερικές εγγραφές(external records)** ενώ αντίστοιχα στο ιδεατό επίπεδο συναντάμε τις έννοιες του **ιδεατού σχήματος (conceptual schema)** και των **ιδεατών εγγραφών (conceptual records)**. Η περιγραφή της ιδεατής εγγραφής περιλαμβάνει πληροφορίες που είναι χρήσιμες για την ορθότητα των δεδομένων, την προστασία τους κ.λπ. και φυσικά το όνομα και το μέγεθος των πεδίων της.

Η **εσωτερική όψη(internal view)** αφορά στην φυσική υλοποίηση της Βάσης Δεδομένων και ορίζει το **εσωτερικό σχήμα(internal schema)** και περιγράφεται με την εσωτερική ή φυσική εγγραφή (**internal record**). Για των ιδεατών και των φυσικών εγγραφών υπάρχουν διαθέσιμες οι αντίστοιχες ιδεατές και φυσικές γλώσσες ορισμού δεδομένων. Παρατηρούμε ότι από τις δύο γλώσσες αυτές, η μια χρησιμοποιείται για τον ορισμό των δεδομένων ενώ η άλλη για την διαχείριση αυτών. Οι γλώσσες ορισμού δεδομένων χρησιμοποιούνται μόνο κατά το σχεδιασμό ή την τροποποίηση του σχεδιασμού μιας εφαρμογής (**Data Base Description, DBD Language**). Η παραδοχή ότι τα δεδομένα παύουν να έχουν νόημα μετά την εκτέλεση ενός προγράμματος κάνει τις συμβατικές γλώσσες να ενοποιούν τους ορισμούς και τις επεξεργασίες των δεδομένων. Αντίθετα

όμως σε ένα ΣΔΒΔ τα δεδομένα εξακολουθούν να υπάρχουν στη Βάση Δεδομένων.

Επίσης παρουσιάζονται δύο απεικονίσεις ή μετασχηματισμοί (**mappings**): η απεικόνιση ιδεατού / εσωτερικού σχήματος οι οποίες στην πραγματικότητα εξασφαλίζουν την ανεξαρτησία δεδομένων. Η απεικόνιση, για παράδειγμα, ιδεατού /εσωτερικού σχήματος και η απεικόνιση του εξωτερικού / ιδεατού σχήματος οι οποίες στην πραγματικότητα εξασφαλίζουν την ανεξαρτησία των δεδομένων. Η απεικόνιση, για παράδειγμα, ιδεατού /εσωτερικού περιγράφει πως μια ιδεατή όψη έχει υλοποιηθεί σε έναν ή περισσότερους δίσκους και κάθε αλλαγή στην φυσική υλοποίηση της Βάσης Δεδομένων έχει σαν αποτέλεσμα να αλλάξει η απεικόνιση αυτή. Αυτό μπορεί να το κάνει ο Διαχειριστής της Βάσης Δεδομένων χωρίς να χρειαστεί καμιά τροποποίηση των προγραμμάτων εφαρμογής.

Ένα κύριο τμήμα της αρχιτεκτονικής, όπως είναι εμφανές, είναι το Σύστημα Διαχείρισης Βάσης Δεδομένων που είναι υπεύθυνο για την ορθή διεκπεραίωση των απαιτήσεων χρήσης. Συγκεκριμένα, το Σύστημα Διαχείρισης Βάσης Δεδομένων πραγματοποιεί τις ακόλουθες διαδικασίες:

- ο Δέχεται την περιγραφή χρήσεων σε κάποια εξωτερική γλώσσα διαχείρισης δεδομένων,
- ο Ελέγχει την ορθότητα της χρήσης,
- ο Συμβουλευεται το εξωτερικό σχήμα, την εξωτερική / ιδεατή απεικόνιση, το ιδεατό σχήμα, την ιδεατή /εσωτερική απεικόνιση και το εσωτερικό σχήμα και,
- ο Εκτελεί τις απαραίτητες ενέργειες για την ικανοποίηση της χρήσης.

Θα πρέπει να αναφέρουμε ότι η έννοια της ανεξαρτησίας των δεδομένων αναφέρεται συχνά σαν φυσική ανεξαρτησία δεδομένων (**physical data independence**) στο Σύστημα Διαχείρισης Βάσης

Δεδομένων και αφορά στη δυνατότητα που έχουν αυτά τα συστήματα να αλλάζουν τις εσωτερικές εγγραφές χωρίς να επηρεάζεται το ιδεατό και το εξωτερικό επίπεδο.

3.3.3 ΔΟΜΕΣ ΒΑΣΕΩΝ ΔΕΔΟΜΕΝΩΝ

Γενικά παρατηρούμε ότι η εξωτερική όψη μιας Βάσης Δεδομένων ανταποκρίνεται στον τρόπο που ένας χρήστης μοντελοποιεί και αντιλαμβάνεται τα δεδομένα του συσχετισμένα μεταξύ τους , ανεξάρτητα από τον τρόπο που αυτά φυλάσσονται φυσικά στα μέσα αποθήκευσης (δίσκους). Υπάρχουν τρεις βασικοί τρόποι δόμησης ή συσχέτισης των δεδομένων για την επεξεργασία της εξωτερικής άποψης της Βάσης Δεδομένων:

- ▼ Ο Σχεσιακός(**Relational**)
- ▼ Ο Ιεραρχικός(**Hierarchical**)
- ▼ Ο Δικτυωτός(**Network**)

Στη συνέχεια θα παραθέσουμε μια ανάλυση των χαρακτηριστικών του κάθε τρόπου δόμησης και έμμεσα των διαφορών τους, μέσω ενός παραδείγματος Βάσης Δεδομένων που αφορά την διαχείριση αποθήκης υλικών μιας επιχείρησης με αντικείμενο το χονδρεμπόριο. Κάνοντας μια σύντομη περιγραφή των δραστηριοτήτων αυτής της επιχείρησης θα αναδείξουμε τις οντότητες και τον τρόπο σύνθεσης μεταξύ τους που υπεισέρχονται στην παραγωγή των απαιτούμενων πληροφοριών:

- Ø Αγορές υλικών που απαιτούν την ύπαρξη και συνεχή ενημέρωση αρχείων προμηθευτών και αρχείου υλικών με παράλληλη γνώση και ενημέρωση της εξέλιξης των τιμών αγοράς.

- Ø Διαχείριση αποθήκης που στηρίζεται κατά βάση στο αρχείο υλικών , αλλά με κυρίαρχο χαρακτηριστικό το κρίσιμο απόθεμα(ελάχιστη ποσότητα αποθέματος για την δυνατότητα ανταπόκρισης στις απαντήσεις των πελατών)
- Ø Πωλήσεις υλικών που απαιτούν την ύπαρξη και συνεχή ενημέρωση αρχείου πελατών, αλλά και του αρχείου υλικών σε συνδυασμό με την ύπαρξη μιας τιμολογιακής πολιτικής επί των πωλήσεων:

Με βάση τα παραπάνω οι οντότητες που αναδεικνύονται μέσα από αυτή τη σύντομη ανάλυση είναι:

- § Προμηθευτές με στοιχεία επωνυμίας, ΑΦΜ , εύρος παραρτημάτων κ.λ.π.
- § Πελάτες με αντίστοιχα στοιχεία με τους προμηθευτές
- § Υλικά με τα απαραίτητα δεδομένα διαχείρισης ,όπως τον κωδικό, την ονομασία, τιμές αγοράς και πώλησης , το κρίσιμο απόθεμα κ.λ.π

Είναι φανερό ότι οι τρεις παραπάνω οντότητες θεωρούνται εκ των « ων ουκ άνευ» για τον σχεδιασμό οποιασδήποτε Βάσης Δεδομένων διαχείρισης αγορών-πωλήσεων και συνιστούν τις λεγόμενες φυσικές οντότητες με την έννοια της ανταπόκρισής τους σε οντότητες που έχουν υλική υπόσταση (πρόσωπα ή πράγματα). Όμως η μόνη θεώρηση αυτών των οντοτήτων δεν αρκεί για την ολοκλήρωση της Βάσης Δεδομένων ,δεδομένου ότι μια προσεκτική παρατήρηση δείχνει ότι αυτές είναι εντελώς ανεξάρτητες μεταξύ τους τουλάχιστον από πλευράς χαρακτηριστικών. Επομένως, επιβάλλεται η θεώρηση λογικών οντοτήτων οι οποίες αφενός θα αποκαταστήσουν τη σύνθεση των πιο πάνω φυσικών οντοτήτων, αφετέρου θα περιγράψουν τις δραστηριότητες και διαδικασίες που αναφέρθηκαν. Οι λογικές αυτές οντότητες θα μπορούσαν να είναι :

- Οι αγορές με τα στοιχεία του τιμολογίου αγοράς και συνδέουσες προμηθευτή με υλικό/ά
- Οι πωλήσεις με τα στοιχεία του τιμολογίου πώλησης και συνδέουσες υλικό /ά με προμηθευτή

Συνοψίζοντας θα μπορούσαμε να συμπεράνουμε ότι ανεξάρτητα από την μορφή των πληροφοριακών συστημάτων και του τρόπου Οργάνωσης Δεδομένων τα απαιτούμενα στοιχεία είναι:

- Προμηθευτών
- Πελατών
- Υλικών
- Αγορών
- Πωλήσεων

Θα πρέπει να τονιστεί ότι οι γλώσσες διαχείρισης δεδομένων που χρησιμοποιούνται από τους χρήστες για εκμετάλλευση της Βάσης Δεδομένων ανταποκρίνονται κατά περίπτωση στον τρόπο δόμησης.

α) Ιεραρχικός Τρόπος Δόμησης

Εδώ περιγράφει την δομή μιας ιεραρχικής βάσης δεδομένων με πέντε πληροφοριακούς τομείς(**segments**), όσοι και τα αρχεία που αναδείχθηκαν από την προηγούμενη ανάλυση. Ο τομέας Υλικά βρίσκεται στη ρίζα, δεδομένου ότι λογικά αποτελεί την κυρίαρχη οντότητα της εφαρμογής (αντικείμενο του χονδρεμπορίου). Κάθε υλικό έχει έναν ή περισσότερους προμηθευτές(**1:N**) έναν ή περισσότερους πελάτες(**1:M**). Με τη σειρά τους κάθε προμηθευτής / πελάτης κάνει διάφορες προμήθειες / αγορές.

β) Δικτυωτός Τρόπος Δόμησης

Εδώ παρουσιάζουμε την ίδια βάση δεδομένων ,από πλευράς ανάλυσης, σε δόμηση δικτύου. Η μέθοδος αυτή χρησιμοποιήθηκε ελάχιστα, δεδομένου ότι εμφανίστηκε σχεδόν παράλληλα με την ιεραρχική. Αν και πλεονεκτεί σε ότι αφορά την επανάληψη των ίδιων δεδομένων, που αποφεύγεται σχεδόν στο απόλυτο, εντούτοις υστερεί σε χρόνους προσπέλασης δεδομένου ότι η παραγωγή σύνθετων πληροφοριών απαιτεί αναζήτηση στοιχείων σε πολλές διαφορετικές εγγραφές.

γ) Σχεσιακός Τρόπος Δόμησης

Η δομή μιας Σχεσιακής βάσης δεδομένων με τρεις φυσικές (προμηθευτές, πελάτες, υλικά) και μια λογική οντότητα(κίνηση) και τα ίδια ακριβώς δεδομένα. Παρατηρούμε ότι αυτή η βάση δεδομένων αποτελείται από τέσσερις πίνακες υπό μορφών αρχείων όπου κάθε γραμμή ονομάζεται **συσχέτιση (relationship)** κάποιων **χαρακτηριστικών (attributes)** που περιγράφουν μια **οντότητα(entity)**. Είναι προφανές ότι ο όρος συσχέτιση αντιστοιχεί στην λογική εγγραφή των συμβατικών αρχείων και το χαρακτηριστικό (**attribute**) στην έννοια του πεδίου.

ΚΕΦΑΛΑΙΟ 4^ο

ΛΗΨΗ ΑΠΟΦΑΣΕΩΝ

ΕΙΣΑΓΩΓΗ

Όπως προαναφέρθηκε, ένα από τα βασικά εργαλεία για τη διοίκηση της επιχείρησης είναι η λήψη έγκαιρων και κοντά στην πραγματικότητα αποφάσεων, οι οποίες βασίζονται σε:

- Προβλέψεις
- Επεξεργασία στοιχείων του παρελθόντος
- Επεξεργασία στοιχείων του παρόντος
- Δειγματοληψίες
- Απόψεις και γνώμες
- Υποκειμενικότητα

Αυτές όμως οι απαιτήσεις συγκρούονται στην πράξη με περιορισμούς όπως:

- Οι εναλλακτικές λύσεις δεν είναι δεδομένες, αλλά πρέπει να εντοπιστούν και ν' απομονωθούν κατόπιν έρευνας.
- Ο σαφής προσδιορισμός των αποτελεσμάτων, τα οποία παράγει κάθε εναλλακτική λύση, είναι δύσκολος.
- Συχνά το περιβάλλον της λήψης απόφασης δεν είναι προσδιορίσιμο.
- Η πιθανή ανεπάρκεια του διοικητικού φορέα, αν ληφθεί υπόψη ότι ο εντοπισμός και η κατανόηση του προβλήματος αποτελεί το πιο αποφασιστικό έργο της διοίκησης.
- Η δυνατότητα πληροφόρησης για τα προϊόντα, τις τιμές και τις στρατηγικές των ανταγωνιστών.
- Η δυνατότητα πληροφόρησης για τους παράγοντες του περιβάλλοντος (ζήτηση προϊόντων, μόδα, επίπεδα τιμών κ.λπ.)

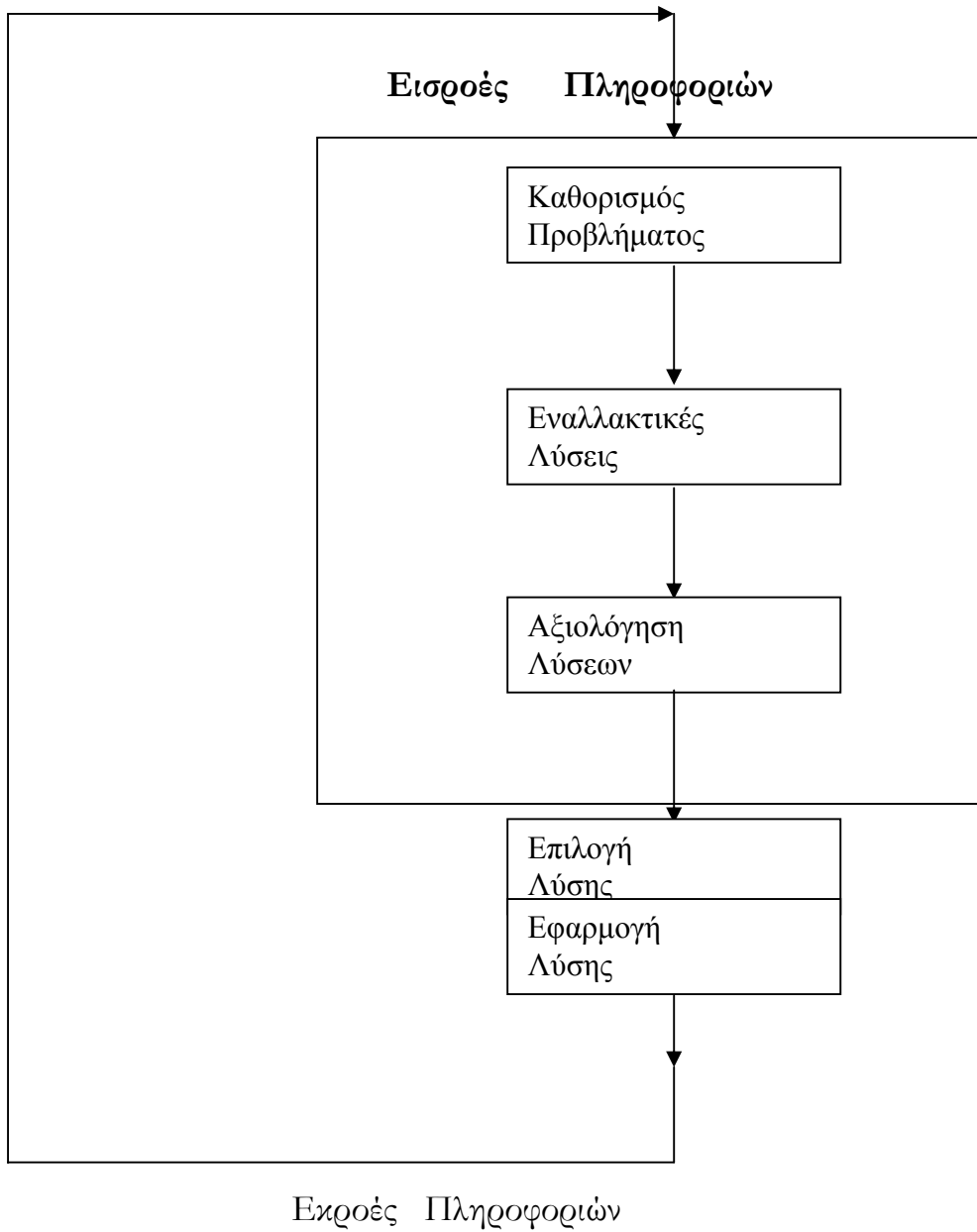
- Η επάρκεια της γενικής πληροφόρησης.
- Η δυνατότητα εντοπισμού και κατανόησης του προβλήματος.
- Η διοικητική πείρα από παρόμοιες περιπτώσεις οι οποίες αντιμετωπίστηκαν στο παρελθόν.
- Η τεχνογνωσία και η νοοτροπία της διοίκησης.
- Το πλήθος των παραγόντων οι οποίοι υπεισέρχονται σε κάθε απόφαση και την ικανοποίηση του στελέχους ν' ασχοληθεί με όλους αυτούς.
- Οι γνώσεις, οι πείρα, το πολιτισμικό επίπεδο και η δημιουργική φαντασία του λαμβάνοντος την απόφαση.
- Ο χρόνος και το χρήμα που μπορεί να διατεθούν για τη διαδικασία λήψης της απόφασης.

B.4.1 ΓΕΝΙΚΗ ΜΕΘΟΔΟΛΟΓΙΑ ΛΗΨΗΣ ΑΠΟΦΑΣΗΣ

Η λήψη αποφάσεων ακολουθεί βασικές λογικές φάσεις. Αποτελεί ένα σύστημα κλειστού βρόχου εισροής , παραγωγής, αξιολόγησης και εκροής πληροφοριών, όπως παρουσιάζεται παραστατικά στο σχήμα **4.1**.

Αρχική Πληροφόρηση

Επαναπληροφόρηση



Σχήμα 4.1: Διαδικασία Λήψης Απόφασης

4.1.1 ΚΑΘΟΡΙΣΜΟΣ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ

Η έναρξη της διαδικασίας λήψης απόφασης γίνεται με την εμφάνιση κάποιου προβλήματος. Ο καθορισμός του προβλήματος απαιτεί τη διατύπωση των πραγματικών ερωτημάτων τα οποία ζητούν απάντηση.

Η ύπαρξη του προβλήματος συνδέεται άμεσα με την ύπαρξη:

- Ø Απόκλισης μεταξύ τεθέντος στόχου και του αποτελέσματος από την υλοποίησή του.
- Ø Απόκλισης μεταξύ νέου στόχου και των απαιτήσεων της τρέχουσας πραγματικότητας.

Το αντικείμενο της μελέτης, σ' αυτήν τη φάση, εστιάζεται στον καθορισμό των επιδιώξεων του λαμβάνοντος την απόφαση. Ο εντοπισμός των πραγματικών ερωτημάτων τα οποία ζητούν απάντηση θα πρέπει να συνοδεύεται από την ύπαρξη προτύπων των στόχων ή των επιθυμητών αποτελεσμάτων προς τα οποία θα κατευθύνονται οι απαντήσεις και με τη βοήθεια των οποίων κρίνεται η σοβαρότητα κάθε προβλήματος (π.χ. αν η εισοδος ενός νέου προϊόντος στην αγορά είναι στρατηγικής σημασίας για την επιχείρηση, η απόκλιση από το στόχο είναι μεγαλύτερης σημασίας από την απόκλιση η οποία παρατηρείται για κάποιο άλλο προϊόν το οποίο δεν παρουσιάζει την ίδια σημασία για την επιχείρηση).

4.1.2 ΕΝΑΛΛΑΚΤΙΚΕΣ ΛΥΣΕΙΣ

Ο καθορισμός των ερωτημάτων τα οποία ζητούν απάντηση ακολουθείται από την παράθεση των διάφορων απαντήσεων οι οποίες μπορεί να ικανοποιούν αυτά τα ερωτήματα. Οι απαντήσεις αυτές αποτελούν και τις εναλλακτικές λύσεις του προβλήματος.

Σ' αυτήν τη φάση παρατίθενται όλες οι πιθανές λύσεις του προβλήματος, καταγράφοντας τις προϋποθέσεις εφαρμογής τους (κόστος-

ωφέλεια), ανεξάρτητα από τη δυνατότητα υλοποίησής τους (η αξιολόγησή τους γίνεται στην επόμενη φάση).

Η αναζήτηση των εναλλακτικών λύσεων στρέφεται στις ακόλουθες κατευθύνσεις:

✓ Να μην υπάρξει καμία αντιμετώπιση του προβλήματος

Αποτελεί τη λύση η οποία προτάσσεται σε κάθε πρόβλημα. Ορισμένες φορές αποτελεί την καλύτερη, οικονομοτεχνικά, λύση, όταν η εμφανισθείσα απόκλιση ή το επιθυμητό αποτέλεσμα δημιουργεί απώλειες μικρότερες από το κόστος κάθε άλλης εναλλακτικής λύσης.

✓ Αναδρομή στο παρελθόν

Εξετάζεται η περίπτωση της εμφανιζόμενης απόκλισης να έχει παρουσιασθεί και να έχει αντιμετωπιστεί επιτυχώς κατά το παρελθόν. Είναι επίσης δυνατό να έχει εμφανισθεί κατά το παρελθόν απόκλιση η οποία, αν και δεν ήταν όμοια με την τρέχουσα, να έχει ωστόσο κοινά σημεία μ' αυτήν. Σε μια τέτοια περίπτωση η πείρα του παρελθόντος μπορεί ν' αποτελέσει έναυσμα για την εξεύρεση λύσεων στην τρέχουσα περίπτωση.

✓ Έλεγχος των παραγόντων της λύσης

Στην περίπτωση κατά την οποία το πρόβλημα είναι κοινοφανές για τον λαμβάνοντα την απόφαση, οι λύσεις θα πρέπει να δημιουργηθούν από το μηδέν. Σ' αυτήν την περίπτωση ο λαμβάνων την απόφαση είναι υποχρεωμένος να εξετάσει τους παράγοντες οι οποίοι υπεισέρχονται στην επίλυση του προβλήματος και να δημιουργήσει ομάδες εναλλακτικών λύσεων με κριτήριο τον έλεγχο αυτών των παραγόντων. Έτσι μπορεί να εμφανιστούν λύσεις των οποίων οι παράγοντες:

- Εξαρτώνται απόλυτα από τον λαμβάνοντα την απόφαση(π.χ. στην περίπτωση ύπαρξης προβλήματος στο λογιστήριο μιας επιχείρησης, ο υπεύθυνος του λογιστηρίου μπορεί να βρεί λύσεις οι οποίες εξαρτώνται αποκλειστικά από παράγοντες οι οποίοι αφορούν μόνο το λογιστήριο, όπως μεταβολή της εργασιακής ροής, αναδιάταξη των εργασιών κ.λπ.).
- Δεν εξαρτώνται διόλου από τον λαμβάνοντα την απόφαση(π.χ. οι λύσεις του προβλήματος του λογιστηρίου απαιτεί την μεταβολή της εργασιακής ροής άλλων τμημάτων της επιχείρησης.).
- Άλλοι εξαρτώνται από τον λαμβάνοντα την απόφαση και άλλοι όχι(π.χ. οι λύσεις του προβλήματος του λογιστηρίου απαιτεί την μεταβολή της εργασιακής ροής και του λογιστηρίου και άλλων τμημάτων της επιχείρησης).

✓ Έλεγχος του χρόνου

Ο χρόνος επηρεάζει τη λήψη απόφασης με δύο τρόπους:

- Σε σχέση με την προσωρινότητα της λύσης. Άλλες λύσεις θεωρούνται μακρόβιες άρα επιζητούν πληρέστερη και σφαιρική μελέτη σε σχέση με ολόκληρη την επιχείρηση(προγραμματισμός), και άλλες προσωρινού ή μεταβατικού χαρακτήρα με αποτέλεσμα να μην απαιτούν μια απόλυτα εμπεριστατωμένη μελέτη.
- Σε σχέση με το χρόνο εφαρμογής της λύσης. Η πίεση χρόνου για την εφαρμογή μιας λύσης, περιορίζει αισθητά την ολοκληρωμένη αναζήτηση εναλλακτικών λύσεων.

4.1.3 ΑΞΙΟΛΟΓΗΣΗ ΛΥΣΕΩΝ

Με την ολοκλήρωση της προηγούμενης φάσης, ο λαμβάνων την απόφαση κατέχει μια ομάδα εναλλακτικών λύσεων οι οποίες θα πρέπει ν' αξιολογηθούν και να ιεραρχηθούν με τη βοήθεια κάποιου κριτηρίου.

Τα βήματα της αξιολόγησης είναι τα ακόλουθα:

- **Διαβάθμιση των χαρακτηριστικών**

Για κάθε λύση καθορίζονται τα πλεονεκτήματα (ωφέλεια, ευκολία εφαρμογής) και τα μειονεκτήματα (κόστος, δυσκολία εφαρμογής), σε σχέση με την τρέχουσα κατάσταση της επιχείρησης.

- **Αποτίμηση του κόστους και της ωφέλειας**

Για κάθε λύση τα πλεονεκτήματα και τα μειονεκτήματα τα οποία παρουσιάζει αποτιμώνται σε κόστος και ωφέλεια για την επιχείρηση.

- **Πρόβλεψη των συνθηκών εφαρμογής**

Η αποτελεσματικότητα εφαρμογής κάθε λύσης εξαρτάται σε μεγάλο βαθμό από τις κρατούσες συνθήκες κατά τη στιγμή αυτής της εφαρμογής (π.χ. η αγορά, η

εγκατάσταση και η λειτουργία ενός Η/Υ στην επιχείρηση εξαρτάται από τις τιμές τις αγοράς κατά τη στιγμή της αγοράς του Η/Υ , από τη διαθεσιμότητα κατάλληλου προσωπικού , από την υπάρχουσα υποδομή στην επιχείρηση κ.λπ.).

Η πρόβλεψη όλων των πιθανών παραγόντων, καθώς και ο βαθμός αξιοπιστίας της πιθανότητας εμφάνισής τους, δημιουργεί για κάθε λύση νέα κλίμακα αξιολόγησης.

4.1.4 ΕΠΙΛΟΓΗ ΤΗΣ ΛΥΣΗΣ

Από τις διάφορες αξιολογημένες λύσεις μια μόνο θα επιλεγεί.

Μια λύση πάντως είναι άμεσα απορριπτέα εφόσον παρουσιάζει κόστος μεγαλύτερο της ωφέλειας.

Μια λύση θεωρείται ότι είναι η καλύτερη (βέλτιστη λύση) όταν:

- Παρουσιάζει το μικρότερο πηλίκον κόστους/ ωφέλειας.
- Το κόστος μπορεί να καλυφθεί από την επιχείρηση.
- Βρίσκεται στην κατεύθυνση του προγραμματισμού της επιχείρησης.
- Γίνεται αποδεκτή από τα άτομα τα οποία θα υλοποιήσουν.
- Ανταποκρίνεται στις χρονικές δεσμεύσεις.
- Παρουσιάζει το μικρότερο κίνδυνο αποτυχίας.
- Παρουσιάζει ευκολία προσαρμογής και αναπροσαρμογής στις κρατούσες συνθήκες της επιχείρησης.

4.2 ΔΙΑΔΙΚΑΣΙΑ ΛΗΨΗΣ ΑΠΟΦΑΣΗΣ

Η γενική μεθοδολογία, που προαναφέρθηκε, αναλύεται στη συνέχεια βήμα προς βήμα.

1) Αντικειμενικοί Στόχοι

Καθορίζονται οι αντικειμενικοί στόχοι. Αυτοί εκφράζουν την επιδίωξη του ατόμου το οποίο λαμβάνει την απόφαση. Βασικοί τέτοιοι στόχοι θεωρούνται:

- Η μεγιστοποίηση του κέρδους ή γενικότερα της ωφέλειας.
- Η ελαχιστοποίηση των απωλειών ή του κόστους.

2) Οι καταστάσεις της Φύσης

Παρατίθενται οι προβλεπόμενες καταστάσεις της φύσης. Πρόκειται για τους διάφορους παράγοντες οι οποίοι επηρεάζουν την εφαρμογή μιας λύσης (συνθήκες εφαρμογής). Δηλαδή τι εναλλακτικές λύσεις, ενέργειες ή προσδοκίες παρεμβαίνουν στην λήψη απόφασης. Για παράδειγμα, το γεγονός ότι μια μετοχή μπορεί να ανατιμηθεί, υποτιμηθεί ή να παραμείνει σταθερή παρουσιάζει τρεις καταστάσεις της φύσης (ανατίμηση, υποτίμηση, σταθερότητα).

3) Οι πιθανότητες Εμφάνισης των Καταστάσεων της Φύσης

Εκτιμούνται οι πιθανότητες εμφάνισης των καταστάσεων της φύσης. Για κάθε κατάσταση της φύσης ποια είναι η πιθανότητα εμφάνισής της; Η πιθανότητα αυτή μπορεί να υπολογίζεται:

- Θεωρητικά.
- Στη βάση της συχνότητας εμφάνισης της ίδιας κατάστασης στο παρελθόν.
- Δειγματοληπτικά.

4) Οι Εναλλακτικές Στρατηγικές

Καθορίζονται οι εναλλακτικές στρατηγικές. Πρόκειται για τις διαφορετικές λύσεις οι οποίες αντιστοιχούν στην εμφάνιση κάθε μιας από τις καταστάσεις της φύσης. Η στρατηγική εκείνη η οποία παρουσιάζει, μετά την αξιολόγηση, τα καλύτερα αποτελέσματα, στη βάση των κριτηρίων επιλογής, ονομάζεται **κυρίαρχη**.

5) Τα Αναμενόμενα Αποτελέσματα

Εκτίμηση των αναμενόμενων αποτελεσμάτων. Για κάθε κατάσταση της φύσης όταν ακολουθείται μια στρατηγική αναμένεται ένα ορισμένο αποτέλεσμα. Έτσι, για παράδειγμα, η πώληση των μετοχών θα έχει σαν αποτέλεσμα X κέρδος ή ζημιά και η μη πώληση θα έχει σαν αποτέλεσμα Y κέρδος ή ζημιά αντίστοιχα.

6) Κριτήρια Επιλογής

Καθορίζονται τα κριτήρια επιλογής της κυρίαρχης στρατηγικής. Αποτελούν τα κριτήρια με τη βοήθεια των οποίων επιλέγεται η καλύτερη στρατηγική. Τα βασικά κριτήρια επιλογής είναι τα ακόλουθα:

- Κριτήριο της αναμενόμενης αξίας(**Expected value**):Με βάση αυτό το κριτήριο επιλέγεται η στρατηγική η οποία παρουσιάζει το καλύτερο αναμενόμενο θετικό αποτέλεσμα.

- Κριτήριο της αναμενόμενης απώλειας ευκαιρίας (**Expected opportunity loss**): Με βάση αυτό το κριτήριο επιλέγεται η στρατηγική η οποία παρουσιάζει τη μικρότερη αναμενόμενη απώλεια.
- Κριτήριο της αναμενόμενης αξίας της τέλει πληροφόρησης (**Expected value of perfect information**): Συχνά εμφανίζεται η ανάγκη να δαπανηθεί κάποιο ποσό ώστε να συγκεντρωθούν περισσότερες πληροφορίες οι οποίες αυξάνουν την πιθανότητα επιλογής της καλύτερης στρατηγικής.
- Κριτήριο της αναμενόμενης χρησιμότητας (**Expected utility**): Χρησιμοποιείται στις περιπτώσεις στις οποίες η εμφάνιση της αναμενόμενης κατάστασης της φύσης επιφέρει μεγάλη ωφέλεια και η μη εμφάνισή της επιφέρει μεγάλη ζημία (αποφάσεις μεγάλου ρίσκου). Σ' αυτές τις περιπτώσεις εμφανίζεται η έννοια της χρησιμότητας η οποία όμως είναι υποκειμενική (πως εκτιμά την κατάσταση ο λαμβάνων την απόφαση).

7) Το Συνολικό Αναμενόμενο Αποτέλεσμα

Εκτιμάται το συνολικό αποτέλεσμα. Το αναμενόμενο αποτέλεσμα από την εφαρμογή μιας στρατηγικής στις διάφορες καταστάσεις της φύσης αντιστοιχούν σε απώλειες ή σε αμοιβές, ανάλογα με το θεωρούμενο κριτήριο επιλογής. Το συνολικό αποτέλεσμα από την εφαρμογή μιας στρατηγικής στο σύνολο των καταστάσεων της φύσης παράγει μια τιμή εκφραζόμενη στις επιδιώξεις εκείνου ο οποίος λαμβάνει την απόφαση.

8) Μήτρα Αποτελεσμάτων

Δημιουργείται η μήτρα αποτελεσμάτων. Όλοι οι προαναφερόμενοι παράγοντες παρατίθενται σε μήτρα αποτελεσμάτων (**payoff matrix**), όπως πίνακας **11.1**, με τη βοήθεια της οποίας μπορεί να επιλεγεί η στρατηγική εκείνη η οποία εξυπηρετεί καλύτερα τις επιδιώξεις εκείνου του οποίου αποφασίζει.

Στην περίπτωση κατά την οποία στόχος είναι η μεγιστοποίηση του κέρδους προκύπτει η **μήτρα κερδών**. Στην περίπτωση κατά την οποία είναι η ελαχιστοποίηση των απωλειών ή του κόστους προκύπτει η **μήτρα απωλειών**.

ΠΙΝΑΚΑΣ 11.1. ΜΗΤΡΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ		
Καταστάσεις Φύσης	K1 K2 K3 K4 K5 K6	ΣΥΝΟΛΙΚΟ ΑΠΟΤΕΛΕΣΜΑ
Πιθανότητες	P1 P2 P3 P4 P5 P6	
Στρατηγικές	Αποτελέσματα	
Σ1	A11 A12 A13 A14 A15 A16	AA1
Σ2	A21 A22 A23 A24 A25 A26	AA2
Σ3	A31 A32 A33 A34 A35 A36	AA3
Σ4	A41 A42 A43 A44 A45 A46	AA4

9) Συνθήκες Λήψης Απόφασης

Βασικό ρόλο στη λήψη της απόφασης παίζει ο βαθμός γνώσης των πιθανοτήτων εμφάνισης των καταστάσεων της φύσης. Ανάλογα με αυτόν το βαθμό γνώσης δημιουργούνται διαφορετικές συνθήκες κάτω από τις οποίες λαμβάνεται η απόφαση.

Ø Συνθήκες Βεβαιότητας

Στην περίπτωση Λήψης Απόφασης σε Συνθήκες Βεβαιότητας (**Decision Making Under Certainty- DMUC**), θεωρείται σχεδόν βέβαιη η εμφάνιση μιας ορισμένης κατάστασης της φύσης. Κατά συνέπεια η μήτρα κερδών θα περιέχει μόνο μια στήλη, η οποία θ' αναφέρεται στην κατάσταση της φύσης η οποία θα εμφανιστεί με βεβαιότητα.

Ø Συνθήκες Κινδύνου

Στην περίπτωση Λήψης Απόφασης σε Συνθήκες Κινδύνου (**Decision Making Under Risk- DMUR**), θεωρούνται γνωστές οι πιθανότητες εμφάνισης των καταστάσεων της φύσης. Η τελική επιλογή της στρατηγικής εξαρτάται από την επιλογή του κριτηρίου αξιολόγησης. Έτσι, για παράδειγμα, ως υποθεθεί ότι λαμβάνων την απόφαση δίνει τις ακόλουθες πιθανότητες για την εμφάνιση των διάφορων καταστάσεων της φύσης:

P1 = 0.3: η πιθανότητα εμφάνισης της **K1**.

P2 = 0.5 : η πιθανότητα εμφάνισης της **K2**.

P3 = 0.2: η πιθανότητα εμφάνισης της **K3**.

P1+P2+P3 = 1.

Απομένει η επιλογή του κριτηρίου αξιολόγησης της καλύτερης στρατηγικής.

Κριτήριο της αναμενόμενης αξίας. Στην περίπτωση επιλογής αυτού του κριτηρίου το συνολικό αναμενόμενο αποτέλεσμα θα είναι το αναμενόμενο κέρδος. Η μήτρα αποτελεσμάτων θα είναι μήτρα κερδών. Επιλέγεται η στρατηγική η οποία δίνει το μεγαλύτερο κέρδος (τολμηρή άποψη).

Κριτήριο της αναμενόμενης απώλειας ευκαιρίας. Στην περίπτωση επιλογής αυτού του κριτηρίου το συνολικό αναμενόμενο αποτέλεσμα θα είναι η αναμενόμενη απώλεια. Η μήτρα αποτελεσμάτων θα είναι μήτρα απωλειών. Επιλέγεται η στρατηγική η οποία δίνει τη μικρότερη απώλεια (συντηρητική άποψη).

Κριτήριο της αναμενόμενης χρησιμότητας. Σ' αυτήν την περίπτωση η επιλογή της στρατηγικής δεν εξαρτάται από την εφαρμογή των προηγούμενων κριτηρίων αλλά από τις αντικειμενικές δυνατότητες (π.χ. οικονομικές) του λαμβάνοντος την απόφαση και την υποκειμενικότητά του (αίσθηση κινδύνου, χαρακτήρας ριψοκίνδυνος ή όχι). Σε κάθε περίπτωση υπεισέρχεται η έννοια της χρησιμότητας της λύσης η οποία έχει αντικειμενικό (δυνατότητες) και υποκειμενικό (υποκειμενισμός) χαρακτήρα. Ένας τολμηρός "παίκτης" θα προτιμούσε μια στρατηγική με μεγάλο κέρδος ενώ ένας συντηρητικότερος "παίκτης" θα προτιμούσε μια άλλη η οποία του παρέχει μικρότερο κέρδος, στην καλύτερη περίπτωση αλλά και τη μικρότερη, συνολικά, απώλεια στη χειρότερη περίπτωση.

Ø Συνθήκες Αβεβαιότητας

Στην περίπτωση Λήψης απόφασης σε Συνθήκες Αβεβαιότητας(**Decision Making Under Uncertainty- DMUU**), οι πιθανότητες εμφάνισης των καταστάσεων της φύσης είναι άγνωστες. Σε μια τέτοια περίπτωση, είτε γίνεται κάποιος προσδιορισμός αυτών των πιθανοτήτων με ορθολογιστικό ή υποκειμενικό τρόπο και στη συνέχεια η λήψη απόφασης ανάγεται στην προηγούμενη περίπτωση (συνθήκες κινδύνου), είτε εφαρμόζονται κάποια κριτήρια τα οποία εκτίθενται στη συνέχεια.

Κριτήριο του Laplace. Σύμφωνα με αυτό το κριτήριο όλες οι εμφανίσεις των καταστάσεων της φύσης είναι ισοπίθανες. Έτσι $p_1=p_2=p_3=1/3=0.33$ και το πρόβλημα ανάγεται στην περίπτωση κινδύνου.

Κριτήριο Minimax-Maximin του Wald. Η εφαρμογή αυτού του κριτηρίου αποτελεί συντηρητική αντιμετώπιση του προβλήματος. Βασίζεται στην επιλογή της στρατηγικής η οποία είναι λιγότερο δυσμενής.

Αν η μήτρα αναφέρεται σε απώλειες τότε:

-Επιλέγεται η μέγιστη απώλεια (**Max**) κάθε στρατηγικής.

-Επιλέγεται εκείνη η στρατηγική η οποία παρουσιάζει το ελάχιστο (**Min**) από τα προηγούμενα μέγιστα.

Κριτήριο του Hurwics .Η εφαρμογή αυτού του κριτηρίου λαμβάνει υπόψη και την καλύτερη αλλά και τη χειρότερη περίπτωση κάθε στρατηγικής. Πρόκειται για έντονα υποκειμενικό κριτήριο διότι απαιτεί τον καθαρισμό ενός δείκτη αισιοδοξίας (**index of optimism**) α για τον λαμβάνοντα την απόφαση.

$\alpha = 1$ σημαίνει πλήρη αισιοδοξία στη λήψη απόφασης(τόλμη για τη διεκδίκηση του μέγιστου κέρδους ανεξάρτητα από τις συνέπειες αποτυχημένης επιλογής).

$\alpha = 0$ σημαίνει πλήρη απαισιοδοξία στη λήψη απόφασης (επιδιώκεται ο περιορισμός των απωλειών ανεξάρτητα από το κέρδος το οποίο θα προέλθει από επιτυχημένη επιλογή).

Οι τιμές μεταξύ **0** και **1** καθορίζουν αντίστοιχα "μείγματα" αισιοδοξίας και απαισιοδοξίας.

Σε περίπτωση κέρδους:

- Επιλέγονται οι καλύτερες περιπτώσεις κάθε στρατηγικής και πολλαπλασιάζονται επί α .
- Επιλέγονται οι χειρότερες περιπτώσεις κάθε στρατηγικής και πολλαπλασιάζονται επί $1-\alpha$.
- Επιλέγεται η στρατηγική η οποία παρουσιάζει το μεγαλύτερο άθροισμα των δύο γινομένων.

Σε περίπτωση απωλειών:

- Επιλέγονται οι χειρότερες περιπτώσεις κάθε στρατηγικής και πολλαπλασιάζονται επί α .
- Επιλέγονται οι καλύτερες περιπτώσεις κάθε στρατηγικής και πολλαπλασιάζονται επί $1-\alpha$.
- Επιλέγεται η στρατηγική η οποία παρουσιάζει το μικρότερο άθροισμα των δύο γινομένων.

Κριτήριο του Savage (ή κριτήριο της "θλίψης"-regret criterion). Πρόκειται για κριτήριο επιλογής με έντονα απαισιόδοξο χαρακτήρα δεδομένου ότι η λήψη απόφασης γίνεται στην κατεύθυνση του περιορισμού της πιθανής απώλειας και όχι στην κατεύθυνση μεγιστοποίησης του κέρδους.

Ως "θλίψη" $\Theta_{i,j}$ ορίζεται η διαφορά του καλύτερου δυνατού αποτελέσματος το οποίο προκύπτει από την εμφάνιση μιας κατάστασης της φύσης, από το δυνατό

αποτέλεσμα το οποίο μπορεί να εμφανισθεί από αυτήν την κατάσταση της φύσης.

4.3 " ΠΑΙΓΝΙΑ"

Η λήψη απόφασης σε συνθήκες αβεβαιότητας , εφόσον αυτή αφορά πρόβλημα ανταγωνισμού, συχνά αντιμετωπίζεται σαν " παίγνιο "(game). Τα " παίγνια" διακρίνονται στις ακόλουθες κατηγορίες:

- " Παίγνια" Ανθρώπου – Φύσης

Στην περίπτωση αυτή ο άνθρωπος ανταγωνίζεται τις αντιξοότητες οι οποίες εμποδίζουν την επίτευξη του στόχου του και οι οποίες προέρχονται από ανεξέλεγκτους, από αυτόν παράγοντες (π.χ. η παρεμπόδιση της ομαλής παραγωγής μιας επιχείρησης από βλάβες του εξοπλισμού, ασθένειες του προσωπικού, κοινωνικές αναστατώσεις κ.λπ.). Σ' αυτήν την περίπτωση ο ανταγωνιστής του λαμβάνοντας την απόφαση δεν είναι "έξυπνος", δηλαδή δεν αντιστρατεύεται συνειδητά τις ανθρώπινες επιδιώξεις. Τα εμπόδια τα οποία παρεμβάλλονται στις ανθρώπινες επιδιώξεις θεωρούνται τυχαία και υπακούουν στους νόμους των πιθανοτήτων.

- " Παίγνια " μεταξύ Ατόμων

Σ' αυτήν την περίπτωση οι ανταγωνιστές είναι " έξυπνοι ", δηλαδή ενεργούν αναλύοντας κάθε ένας τις ενέργειες του άλλου, προσπαθώντας ν' αποκομίσουν

όφελος , ο ένας σε βάρος του άλλου (π.χ. ανταγωνίστριες εταιρίες οι οποίες παράγουν προϊόντα τα οποία απευθύνονται στην ίδια αγορά).

Στην κατηγορία αυτή το " παίγνιο " έχει τις ακόλουθες ιδιότητες:

- Υπάρχει πεπερασμένος αριθμός ανταγωνιστών.
- Κάθε ανταγωνιστής διαθέτει πεπερασμένο αριθμό εναλλακτικών ενεργειών ο οποίος διαφέρει για κάθε ανταγωνιστή.
- Το " παίγνιο " ολοκληρώνεται όταν κάθε ανταγωνιστής υλοποιήσει τις ενέργειές του.
- Το αποτέλεσμα (αμοιβή) για κάθε παίκτη αποτελεί συνέπεια των επιλεγμένων ενεργειών του συνόλου των ανταγωνιστών και μπορεί να είναι αρνητικό(απώλεια), μηδέν ή θετικό(κέρδος).
- Κάθε ανταγωνιστής διατάσσει τις εναλλακτικές ενέργειές του σε στρατηγικές. Η καθαρή (**pure**) στρατηγική επιλέγεται πριν από την έναρξη του " παιχνιδιού" και δεν λαμβάνει υπόψη τις επιλογές των άλλων ανταγωνιστών. Η μεικτή (**mixed**) στρατηγική περιλαμβάνει προεπιλεγμένους τρόπους ενεργειών αλλά και εναλλακτικές ενέργειες οι οποίες θα πρέπει να γίνουν όταν εκδηλώνονται οι προθέσεις των άλλων ανταγωνιστών.

Ουσιαστικά κάθε ανταγωνιστής προσπαθεί να μαντέψει τις κινήσεις των άλλων ανταγωνιστών του και να προσαρμόσει σ' αυτές τις δικές του ενέργειες.

Αυτή η κατηγορία " παιχνιδιών " διαιρείται σε δύο υποκατηγορίες :

- “ Παιγνια ” με συνολικό αποτέλεσμα μηδέν (**zero-sum games**). Το αλγεβρικό άθροισμα των απωλειών και των κερδών των ανταγωνιστών είναι μηδέν. Δηλαδή όσα χάνουν κάποιοι ανταγωνιστές τα κερδίζουν οι υπόλοιποι(π.χ. αν κάποιος πουλήσει τις μετοχές του και αυτές οι μετοχές αυξηθούν τότε ο πωλητής ζημιώνεται τόσο όσο κερδίζουν οι αγοραστές τους).
- “ Παιγνια ” με συνολικό αποτέλεσμα διάφορο του μηδενός(**non zero-sum games**). Σ’ αυτήν την περίπτωση παράγοντες εκτός των ανταγωνιστών παράγουν κέρδη ή ζημιές τα οποία διαφοροποιούν το συνολικό αποτέλεσμα(π.χ. κίνητρα ή αντικίνητρα από το κράτος σε κάποια κατηγορία ανταγωνιστικών εταιριών στον ανταγωνισμό τους στην αγορά).

4.4 ΤΥΠΟΙ ΚΑΙ ΚΑΤΗΓΟΡΙΕΣ ΑΠΟΦΑΣΕΩΝ

Ο τρόπος με τον οποίο λαμβάνεται μια απόφαση δεν είναι καθορισμένος. Εξαρτάται από πολλούς παράγοντες, όπως ο τύπος της απόφασης, η υποκειμενικότητα του ατόμου το οποίο λαμβάνει την απόφαση, η οργανωτική αντίληψη του επιχειρηματικού περιβάλλοντος.

4.4.1 ΤΥΠΟΙ ΑΠΟΦΑΣΕΩΝ

H.A.SIMON:	P.DRUCKER
<ul style="list-style-type: none"> • Προγραμματισμένες (programmed), οι οποίες αφορούν Επαναλαμβανόμενα προβλήματα , έχουν συγκεκριμένη δομή και Σχεδόν τυποποιημένο τρόπο Αντιμετώπισης. • Απρογραμμάτιστες (non Programmed), οι οποίες αναφέρονται Σε πρωτοεμφανιζόμενα προβλήματα, Αδόμητα για τα οποία δεν υπάρχει Συγκεκριμένος τρόπος αντιμετώπισης. 	<ul style="list-style-type: none"> * Τακτικές(tactical), οι οποίες στρέφονται στη φάση της υλοποίησης του προγραμματισμού, με κύριους αντικειμενικούς στόχους την μεγιστοποίηση του Κέρδους ή την ελαχιστοποίηση του κόστους. * Στρατηγικές (strategic), οι οποίες αναφέρονται στον καθορισμό αντικειμενικών στόχων, τον Προσδιορισμό των διαθέσιμων Πόρων κ.λπ.

A .I. DEBBECK	E. Καρασαββίδου
<ul style="list-style-type: none"> • Συνηθισμένες(routine), οι οποίες παρουσιάζουν μεγάλη συχνότητα εμφάνισης. • Δημιουργικές(creative), οι οποίες αφορούν πρωτοεμφανιζόμενα προβλήματα. • Διαπραγμάτευσης(negotiated), οι οποίες αποτελούν αντικείμενο διαπραγμάτευσης. 	<p>* Αποφάσεις προγραμματισμένες, Συνηθισμένες, τακτικές και διαπραγμάτευσης, οι οποίες λαμβάνονται συνήθως κάτω από Συνθήκες βεβαιότητας, αφορούν μεσαία και κατώτερα διοικητικά στελέχη και αφορούν διαχείριση Αποθεμάτων, κατανομή πόρων, διανομή προϊόντων κ.λπ.</p> <p>* Αποφάσεις απρογραμμάτιστες, Στρατηγικές, δημιουργικές, και Καινοτομίας, οι οποίες λαμβάνονται Συνήθως κάτω από συνθήκες Αβεβαιότητας ή κινδύνου, αφορούν τα Ανώτερα στελέχη και σχετίζονται με τον καθορισμό γενικών στόχων της επιχείρησης, την προσαρμογή της πολιτικής της στις συνθήκες του περιβάλλοντος κ.λπ.</p>

Ο τρόπος με τον οποίο ο λαμβάνων τις αποφάσεις προσεγγίζει την αντιμετώπιση κάθε προβλήματος χαρακτηρίζεται από την υποκειμενικότητα του ατόμου, δηλαδή πως αντιλαμβάνεται την αντικειμενική πραγματικότητα.

Σε γενικές γραμμές ακολουθούνται οι ακόλουθες προσεγγίσεις:

- **Με κριτήριο την ορθολογικότητα (Rationality)**

Ο λαμβάνων την απόφαση θεωρεί ότι υπάρχουν δεδομένες οι καταστάσεις της φύσης και σε κάθε δυνατό συνδυασμό τους αντιστοιχείται ένα αναμενόμενο αποτέλεσμα. Επιλέγεται εκείνη η λύση η οποία θεωρείται ως η πλέον συμφέρουσα. Αυτό η προσέγγιση μπορεί να χρησιμοποιηθεί εφόσον υπάρχει περιορισμένος αριθμός καταστάσεων της φύσης οι οποίες είναι σχετικά γνωστές (βεβαιότητα). Στην περίπτωση όμως κατά την οποία οι καταστάσεις της φύσης είναι πολλές ή απρόβλεπτες αυτή η προσέγγιση είναι πραγματικά ανεφάρμοστη. Για παράδειγμα, ποια θα είναι τα συναισθήματα κάποιου πελάτη σε μια διαδικασία πώλησης; Μπορεί ο πωλητής με βεβαιότητα να καθορίσει την τακτική του απέναντι σε κάθε πελάτη;

- **Με κριτήριο την περιορισμένη ορθολογικότητα και την ικανοποίηση(Bounded rationality and satisficing)**

Σύμφωνα μ' αυτήν την προσέγγιση ο λαμβάνων την απόφαση κινείται με κριτήριο τις γνώσεις του, την εμπειρία του και την ικανοποίηση την οποία αισθάνεται από κάποια αποτελέσματα συνδυασμών των καταστάσεων της φύσης. Η τελική επιλογή του δεν καθορίζεται ανάμεσα από όλες τις δυνατές καταστάσεις της φύσης αλλά ανάμεσα σε εκείνες τις οποίες κατανοεί καλύτερα, αισθάνεται ότι τις ελέγχει και τον οδηγούν προς το αποτέλεσμα το οποίο τον ικανοποιεί περισσότερο και όχι.

- **Αυξητική λήψη απόφασης(Incremental decision making)**

Σ' αυτήν την περίπτωση ο λαμβάνων την απόφαση προσανατολίζεται στην επιλογή λύσης η οποία αυξάνει την ωφέλεια προηγούμενης απόφασης. Ουσιαστικά η επιλογή λύσεων σχετίζεται άμεσα με προηγούμενες λύσεις, των οποίων βελτιώνουν το αποτέλεσμα.

4.4.2 ΟΡΓΑΝΩΤΙΚΟ ΠΛΑΙΣΙΟ ΚΑΙ ΑΠΟΦΑΣΕΙΣ

Εκτός από την υποκειμενική προσέγγιση της λήψης απόφασης, σημαντικό ρόλο παίζει και το οργανωτικό περιβάλλον της επιχείρησης το οποίο καθορίζει τη δομή και τα χαρακτηριστικά των πολιτικών της επιχείρησης.

- **Γραφειοκρατικά μοντέλα λήψης απόφασης
(Bureaucratic models of decision making)**

Αυτά τα μοντέλα λαμβάνουν υπόψη την διαμορφωμένη από το παρελθόν οργανωτική δομή της επιχείρησης. Έχει διαμορφωθεί μια τυποποιημένη οργανωτική δομή, στην οποία οι διαδικασίες και εργασίες είναι προκαθορισμένες και σε μεγάλο βαθμό τυποποιημένες. Η λειτουργία κάθε ατόμου, στα πλαίσια μιας τέτοιας επιχείρησης είναι προκαθορισμένη (λόγω προδιαγραφών ή λόγω συνήθειας) και τυποποιημένη. Αυτό έχει σαν συνέπεια μια λήψη απόφασης:

- Να κατακερματίζεται σε διάφορα οργανωτικά επίπεδα.
- Ν' ακολουθεί προκαθορισμένες προδιαγραφές σε σχέση με τη διαδικασία λήψης της.

Κατά συνέπεια ο ρόλος της υποκειμενικότητας του λαμβάνοντος την απόφαση περιορίζεται σημαντικά. Η τυποποίηση περιορίζει την ένταση της υποκειμενικότητας.

- **Πολιτικά μοντέλα λήψης απόφασης(Political models of decision making)**

Σ' αυτήν την περίπτωση λαμβάνεται υπόψη η συγκέντρωση εξουσίας στα διάφορα οργανωτικά επίπεδα της επιχείρησης και οι συγκρούσεις τις οποίες αυτή συνεπάγεται.

Η λήψη απόφασης περιορίζεται από την επιδιωκόμενη "ισορροπία" μεταξύ των διάφορων κέντρων εξουσίας.

- **Μοντέλα "απορριμάτων" ("Carbage can" models)**

Αυτά τα μοντέλα βασίζονται στην άποψη ότι:

- Η επιχείρηση δεν μπορεί να έχει ορθολογική βάση(δεν μπορούν να προβλεφθούν όλες οι καταστάσεις της φύσης σε κάθε περίπτωση-"χαωτική" αντιμετώπιση).
- Η λήψη απόφασης είναι ένα "τυχαίο" γεγονός στο οποίο εμπλέκονται τυχαία τα προβλήματα, οι καταστάσεις της φύσης και οι λύσεις. Συνεπώς είναι τυχαίο και το αποτέλεσμα της εφαρμογής μιας επιλεγμένης λύσης.
- Η μεθοδολογική και συστηματική αναζήτηση λύσεων είναι άσκοπη(για τα "σκουπίδια") διότι οδηγούν τις επιχειρήσεις να εφευρίσκουν προβλήματα και τα διοικητικά στελέχη να δικαιολογούν το μισθό τους.

- Κάθε πρόβλημα αντιμετωπίζεται κατά τη στιγμή της εμφάνισής του με βάση τα δεδομένα εκείνης της στιγμής.

Η άποψη αυτή, η οποία απορρίπτει εξολοκλήρου την μεθολογική λήψη αποφάσεων, βασίζεται στη διαπίστωση ότι παρά την ύπαρξη εξειδικευμένων συστημάτων λήψης αποφάσεων, μεγάλος αριθμός επιχειρήσεων δεν επιβιώνει στον ανταγωνισμό.

4.5 ΜΟΝΤΕΛΑ ΑΠΟΦΑΣΕΩΝ

Με βάση τα προαναφερόμενα είναι προφανές ότι η λήψη απόφασης για να οδηγήσει σε επιλογή και εφαρμογή μιας λύσης θα πρέπει να πάρει συγκεκριμένη μορφή, η οποία να λαμβάνει υπόψη όλες τις προαναφερόμενες παραμέτρους. Η συγκεκριμενοποίηση και η τυποποίηση (μερική ή ολική) της διαδικασίας λήψης της απόφασης οδηγεί στο **Μοντέλο Απόφασης (Decision Model)**. Η απόλυτη τυποποίηση οδηγεί στις **Δομημένες Αποφάσεις**. Έτσι, για παράδειγμα, αν η επιλογή των νέων υπαλλήλων στην επιχείρηση ακολουθεί μία συγκεκριμένη μεθοδολογία (προδιαγραφή, θέσης εργασίας, καθορισμός προσόντων υποψηφίου, πως ακριβώς γίνεται η συνέντευξη, πως ακριβώς γίνεται η αξιολόγηση κ.λπ.), τότε η επιλογή ενός νέου υπαλλήλου αποτελεί ένα μοντέλο απόφασης. Ο καθορισμός του μοντέλου βασίζεται συχνότατα στα μαθηματικά, την επιχειρησιακή έρευνα, τη στατιστική και τη θεωρία πιθανοτήτων.

Η χρήση των μοντέλων αποφάσεων παρουσιάζει τα ακόλουθα πλεονεκτήματα:

- Ορισμένες φορές τυποποιούνται και χρησιμοποιούνται γενικευμένα από πλήθος διοικητικών στελεχών.
- Επιταχύνουν τη λήψη απόφασης.
- Περιορίζουν, σε λογικά πλαίσια, την υποκειμενικότητα του λαμβάνοντος την απόφαση.

- Λαμβάνονται παρόμοιες αποφάσεις κάτω από παρόμοιες συνθήκες.

Κατηγοροποίηση Μοντέλων

Τα μοντέλα αποφάσεων μπορεί να κατηγοροποιηθούν χρησιμοποιώντας διαφορετικά κριτήρια.

Ως προς την δυνατότητα γενίκευσης την οποία παρέχουν διακρίνονται σε:

- **Εικονικά(iconic):** Τα χαρακτηριστικά του συστήματος παριστάνονται υπό κλίμακα(μακέτες, σχέδια, φωτογραφίες κ.λπ.)
- **Αναλογικά(Analogue):**Τα χαρακτηριστικά του συστήματος παριστάνονται από άλλα (διάγραμμα ροής πληροφοριών, σχέδια ηλεκτρικών εγκαταστάσεων κ.λπ.).
- **Συμβολικά(Symbolic):**Χρησιμοποιούν σύμβολα για την αναπαράσταση του συστήματος(αλγόριθμοι, μαθηματικές σχέσεις κ.λπ.).

Ως προς τον έλεγχο των παραμέτρων οι οποίες υπεισέρχονται στα μοντέλα, αυτά διακρίνονται σε:

- **Περιγραφικά (descriptive):** Δεν περιέχουν ελεγχόμενες παραμέτρους και η λειτουργία του συστήματος περιγράφεται λεκτικά.
- **Επεξηγηματικά(explanatory):** Όλες οι περιεχόμενες παράμετροι είναι ελεγχόμενες και η περιγραφή του συστήματος γίνεται με την παράθεση των σχέσεων μεταξύ των παραμέτρων(αλγόριθμοι με τη βοήθεια της στατιστικής, αριθμητικής ανάλυσης κ.λπ.).

Ως προς την εξέλιξή τους στο χρόνο τα μοντέλα διακρίνονται σε:

- **Στατιστικά(static):** Παρέχουν λύσεις μόνο κάποια χρονική στιγμή.
- **Δυναμικά(dynamic):** Παρέχουν λύσεις κάθε χρονική στιγμή.

Ως προς τον τύπο και το περιεχόμενό τους τα μοντέλα διακρίνονται σε:

- **Αναμονής(queueing):** Αναφέρεται στο σειριακό ρυθμό εξυπηρέτησης στοιχείων τα οποία φθάνουν σε σταθμό εξυπηρέτησης (ουρές αναμονής σε ταμεία, επεξεργασία προγραμμάτων από Η/Υ κ.λπ.).
- **Αποθεμάτων(inventory):** παρέχουν τρόπους διαχείρισης των αποθεμάτων
- **Κατανομής(allocation):** Παρέχουν τρόπους κατανομής πόρων σε αντίστοιχα έργα.
- **Αντικατάστασης και συντήρησης(replacement and maintenance):** Επιτρέπουν την παρακολούθηση της φθοράς του εξοπλισμού και προβλέπουν την απαιτούμενη προληπτική αντικατάσταση εξαρτημάτων καθώς και την προληπτική συντήρηση του εξοπλισμού.
- **Δρομολογιών(routing and scheduling):** Επιτρέπουν τη δρομολόγηση των διαδικασιών ενός έργου(μέθοδοι **PERT**, **COST** κ.λπ.). Π.χ. χρονικός προγραμματισμός έργου, βέλτιστα δρομολόγια διανομής εμπορευμάτων κ.λπ.
- **Αναζήτησης(search):** Επιτρέπουν την αναζήτηση των προϋποθέσεων ώστε να μεγιστοποιείται η πιθανότητα επίτευξης ενός στόχου(π.χ. μεγιστοποίηση κέρδους).
- **Ανταγωνισμού(competition):** Αποτελούν μοντελοποίηση των "παιγνίων".

Ως προς τον τρόπο επεξεργασίας τους από Η/Υ τα μοντέλα διακρίνονται σε:

- **Αλγοριθμικά (algorithmic):** Η λύση παρέχεται μέσω μαθηματικής σχέσης των παραμέτρων του μοντέλου.
- **Ευρετικά(heuristic):** Η απόφαση λαμβάνεται με τη βοήθεια εμπειρικών ή λογικών κανόνων(π.χ. προσομοίωση).

Ως προς τη συμβολή τους στη διοίκηση τα μοντέλα διακρίνονται σε:

- Βελτιστοποίησης(optimization):** Επιτρέπουν τον εντοπισμό της βέλτιστης μεταξύ εναλλακτικών λύσεων.
- **Προσομοίωσης(simulation):** Επιτρέπουν την πρόβλεψη της εξέλιξης στο μέλλον μιας σημερινής κατάστασης υπό προϋποθέσεις.

Β.4.6 ΛΗΨΗ ΑΠΟΦΑΣΕΩΝ ΚΑΙ ΠΛΗΡΟΦΟΡΙΑΚΑ ΣΥΣΤΗΜΑΤΑ

Η λήψη των αποφάσεων σχετίζεται άμεσα με τα πληροφοριακά συστήματα εφόσον και τα δύο αποτελούν εργαλεία διοίκησης – διαχείρισης. Έτσι στην παρούσα παράγραφο γίνεται αναφορά στη σχέση λήψης απόφασης με τα πληροφοριακά συστήματα εξεταζόμενη από διαφορετικές οπτικές γωνίες.

Β.4.6.1 ΔΙΑΔΙΚΑΣΙΑ ΛΗΨΗΣ ΑΠΟΦΑΣΗΣ ΚΑΙ ΠΛΗΡΟΦΟΡΙΑΚΑ ΣΥΣΤΗΜΑΤΑ

Με βάση το σχήμα 11.1 μπορεί να εξεταστεί η συνδρομή των πληροφοριακών συστημάτων στα διάφορα βήματα της διαδικασίας λήψης απόφασης.

Στο βήμα του καθορισμού του προβλήματος κυρίαρχο ρόλο παίζει η αντίληψη ότι κάτι συμβαίνει κάπου στην επιχείρηση, για συγκεκριμένους λόγους και με κάποιες συνέπειες στη λειτουργία της επιχείρησης. Αυτή η αντίληψη αποτελεί προϊόν του διοικητικού και λειτουργικού ελέγχου της επιχείρησης. Κύρια πηγή πληροφοριών μπορεί ν' αποτελεί ένα **MIS**.

Στο βήμα της καταγραφής των εναλλακτικών λύσεων απαιτείται μια πιο προσεκτική θεώρηση της λειτουργίας της επιχείρησης. Σ' αυτήν την περίπτωση μπορεί να αξιοποιηθεί κάποιο μικρό **DSS**.

Στα βήματα της αξιολόγησης των εναλλακτικών λύσεων και της επιλογής της τελικής λύσης απαιτούνται τεχνικές οι οποίες επιτρέπουν την αξιολόγηση των συνεπειών στην επιχείρηση από την επιλογή κάθε εναλλακτικής λύσης. Ένα μεγάλο **DSS** μπορεί ν' αποτελέσει χρήσιμο εργαλείο για τον λαμβάνοντα την απόφαση.

Στο βήμα της εφαρμογής της επιλεγμένης λύσης βασικό ζητούμενο είναι η αξιολόγηση του αποτελέσματος της εφαρμοσμένης λύσης. Δεδομένου ότι αυτό αποτελεί προϊόν ελέγχου, το **MIS** μπορεί ν' αποτελέσει και πάλι χρήσιμο εργαλείο.

B.4.6.2 ΟΡΓΑΝΩΣΗ ΚΑΙ ΠΛΗΡΟΦΟΡΙΑΚΑ ΣΥΣΤΗΜΑΤΑ

Στο στρατηγικό οργανωτικό επίπεδο η λήψη αποφάσεων αφορά τον καθορισμό των στόχων, των πόρων και των πολιτικών της επιχείρησης. Το μεγαλύτερο πρόβλημα αυτών των αποφάσεων είναι η απαίτηση για μελλοντικές προβλέψεις σε σχέση με την εξέλιξη της επιχείρησης και του περιβάλλοντός της. Πρόκειται για αποφάσεις αδόμητες, απρογραμμάτιστες, στρατηγικές,

δημιουργικές και καινοτομίας, οι οποίες λαμβάνονται κάτω από συνθήκες αβεβαιότητας ή κινδύνου. Η λήψη αυτών των αποφάσεων μπορεί να υποστηριχθεί από κάποιο **ESS**.

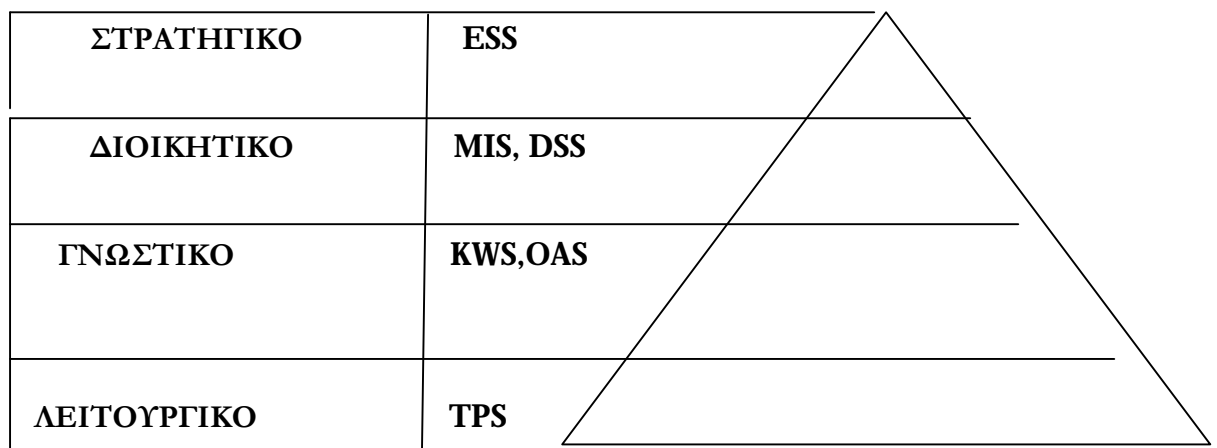
Στο λειτουργικό οργανωτικό επίπεδο αντίθετα οι λαμβανόμενες αποφάσεις είναι δομημένες, προγραμματισμένες, συνηθισμένες και τακτικές οι οποίες αφορούν το λειτουργικό έλεγχο(πόσο τ' αποτελέσματα των εργασιών ανταποκρίνονται στις προδιαγραφές τους) και βασικό ρόλο για τη λήψη τους παίζει η πληροφόρηση η οποία προέρχεται από τα **TPS**.

Στα δύο άλλα οργανωτικά επίπεδα (διοικητικό, γνωστικό) οι αποφάσεις είναι ημι-δομημένες, προγραμματισμένες ή όχι , συνηθισμένες ή όχι, τακτικές. Αναλυτικότερα:

- Στο διοικητικό οργανωτικό επίπεδο οι αποφάσεις αφορούν κυρίως το διοικητικό έλεγχο(πόσο αποτελεσματική και ικανή είναι η λειτουργία της επιχείρησης) και η βασική πληροφόρηση μπορεί να προέρχεται από τα **MIS** και **DSS**.
- Στο γνωστικό οργανωτικό επίπεδο οι αποφάσεις αφορούν την παραγωγή νέων ιδεών για προϊόντα και υπηρεσίες και τη διάδοσή τους μέσα στην επιχείρηση. Εργαλεία για τη λήψη αυτών των αποφάσεων μπορεί ν' αποτελούν τα **KWS** και **OAS**.

Το σχήμα **4.2** παρουσιάζει τους διάφορους τύπους συστημάτων σε σχέση με την οργανωτική πυραμίδα της επιχείρησης.

ΕΠΙΠΕΔΑ ΟΡΓΑΝΩΣΗΣ ΤΥΠΟΣ ΣΥΣΤΗΜΑΤΟΣ



Σχ.4.2 Τύποι Πληροφοριακών Συστημάτων και Οργανωτική Πυραμίδα.

Β.4.6.3 ΑΠΟΦΑΣΕΙΣ ΚΑΙ ΣΧΕΔΙΑΣΜΟΣ ΠΛΗΡΟΦΟΡΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

Οι απαιτήσεις των διοικητικών στελεχών από τα πληροφοριακά συστήματα είναι συνάρτηση των αυξημένων αναγκών τους σε πληροφόρηση μέσα σ' ένα σύνθετο και ραγδαία μεταβαλλόμενο περιβάλλον.

Τα διοικητικά στελέχη απαιτούν:

- Τυποποιημένα συστήματα για το σχεδιασμό, την οργάνωση και τον συντονισμό.
- Εύκολα μεταβαλλόμενα και πολυδύναμα συστήματα τα οποία επιτρέπουν την κάλυψη ευρέος φάσματος απαιτήσεων και την προσαρμογή τους στις μεταβολές του περιβάλλοντος.

Όμως αν και τα τυποποιημένα συστήματα έχουν μεγάλες δυνατότητες στο λειτουργικό τομέα της επιχείρησης, διαθέτουν περιορισμένες δυνατότητες στο διοικητικό τομέα δεδομένου ότι η διοίκηση περιέχει μεγάλο ποσοστό ατυποποίητης ανθρώπινης αυτενέργειας(διαίσθηση, ψυχολογική συγκρότηση κ.λπ.). Σ' αυτό θα πρέπει να προστεθεί ότι μια απόφαση διαφέρει από άτομο σε άτομο και από κατάσταση σε κατάσταση ως προς τη σημασία των στόχων, την ποιότητα του λαμβάνοντος την απόφαση, το οργανωτικό μοντέλο κ.λπ.

Τα προαναφερόμενα έχουν άμεσες επιπτώσεις στο σχεδιασμό νέων συστημάτων. Διότι η λήψη αποφάσεων σε μια επιχείρηση λαμβάνεται από ομάδες ατόμων με τη χρήση οργανωτικών διαδικασιών και τα συστήματα καλούνται να υποστηρίξουν και τις ομάδες αυτές αλλά και τις οργανωτικές διαδικασίες.

Λαμβάνοντας υπόψη τα προηγουμένως γραφόμενα, τα βασικότερα χαρακτηριστικά των συστημάτων, τα οποία θα πρέπει να σεβαστεί ο σχεδιαστής νέων συστημάτων θα πρέπει να είναι:

- Η ευλυγισία τους, προσφέροντας στους χρήστες τους τις περισσότερες δυνατές εναλλακτικές δυνατότητες στη συλλογή και αξιολόγηση των δεδομένων και στις μεταβολές των περιβαλλοντικών συνθηκών.
- Η δυνατότητα υποστήριξης ατόμων με διαφορετικές απόψεις διοίκησης, με διαφορετικές ικανότητες, με διαφορετικές γνώσεις, με τον

υποκειμενισμό τους, τα οποία λειτουργούν σε διαφορετικά οργανωτικά περιβάλλοντα.

- Η δυνατότητα παροχής μοντέλων με ισχυρές αναλυτικές και συνθετικές ικανότητες τα οποία επιτρέπουν την παρουσίαση εναλλακτικών αποτελεσμάτων με βάση τα εισερχόμενα δεδομένα.
- Ν' ανταναικλούν τις οργανωτικές και πολιτικές αναζητήσεις της επιχείρησης.
- Ν' ακολουθούν τις μεταβολές της οργάνωσης, της πολιτικής και των διαδικασιών και να είναι γνωστές οι δυνατότητές τους (τι μπορούν να κάνουν και τι όχι).

ΚΕΦΑΛΑΙΟ 5^ο

ΟΙ ΤΕΧΝΙΚΕΣ ΕΞΟΡΥΞΗΣ ΑΠΟ ΔΕΔΟΜΕΝΑ

5.1 ΕΙΣΑΓΩΓΗ

Οι τεχνικές εξόρυξης από δεδομένα αποτελούν ένα από τα σημαντικότερα εργαλεία ανάλυσης δεδομένων των τελευταίων δεκαετιών . Όλο και περισσότερο αναγνωρίζεται η συνεισφορά των δια φόρων τεχνικών αυτών στην ανάλυση μεγάλων βάσεων δεδομένων σε σχέση με τις κλασικές στατιστικές μεθόδους .

Έτσι, ενώ τα στατιστικά τεστ απαιτούν έλεγχο στατιστικών υποθέσεων σε σχέση με τη μορφή κανόνων σε μεγάλες βάσεις δεδομένων (**Aha, D. , 1997**).

Κατά συνέπεια, η αποτελεσματικότητα των κλασικών μεθόδων περιορίζεται από ένα μεγάλο αριθμό παραγόντων , όπως η ικανότητα του αναλυτή να θέσει τα καίρια ερωτήματα και να έχει άμεσα αποτελέσματα , η δυσκολία στη διαχείριση της πολυπλοκότητας των χαρακτηριστικών κ.ο.κ.

Σε αντίθεση με αυτές τις μεθόδους , οι τεχνικές εξόρυξης από τα δεδομένα επιτρέπουν στον χρήστη να ανακαλύψει τις σημαντικές σχέσεις μέσα στα δεδομένα , χρησιμοποιώντας μεθοδολογίες ανεύρεσης γνώσης (**knowledge discovery**).

Πολλές τεχνικές που χρησιμοποιούνται σήμερα ως τεχνικές εξόρυξης από δεδομένα , υπάρχουν πολλά χρόνια , μέσα στα πλαίσια της επιστήμης της τεχνητής νοημοσύνης (**artificial intelligence**) και αναπτύχθηκαν στα τέλη της δεκαετίας του **1980**.

Παρ' όλο που υπάρχουν διαφορές σε σχέση με τον ακριβή τρόπο ορισμό των τεχνικών εξόρυξης από δεδομένα , οι περισσότεροι επιστήμονες και αναλυτές συμφωνούν ότι οι τεχνικές εξόρυξης από δεδομένα είναι ένα πολυεπιστημονικό

πεδίο γνώσης , το οποίο δανείζεται ιδέες από τα πεδία της επιστήμης της πληροφορικής , όπως την τεχνητή νοημοσύνη και την αναγνώριση προτύπων, τη στατιστική, τη μαθηματική βελτιστοποίηση κ.α. Σύμφωνα με έναν ευρέως αποδεκτό ορισμό, οι τεχνικές εξόρυξης από δεδομένα είναι μια διαδικασία , η οποία ασχολείται με την αναγνώριση προτύπων , σχέσεων, ανωμαλιών και στατιστικός σημαντικών δομών σε δεδομένα.

Οι τεχνικές εξόρυξης από δεδομένα περιλαμβάνουν δυο βασικά βήματα : **1)** την προεπεξεργασία δεδομένων και **2)** την αναγνώριση προτύπων στα δεδομένα αυτά . Κατά το στάδιο της προεπεξεργασίας των δεδομένων αναγνωρίζονται τα σχετικά με το πρόβλημα χαρακτηριστικά (**attributes/ features**). Το δεύτερο στάδιο περιλαμβάνει αναγνώριση προτύπων βάση των χαρακτηριστικών που είχαν αναγνωριστεί στο προηγούμενο στάδιο και την αξιολόγησή τους . Παρ' όλο που κατά τη διαδικασία εξόρυξης από δεδομένα δίνεται ιδιαίτερη έμφαση στο στάδιο της αναγνώρισης προτύπων , αξίζει να σημειωθεί ότι πολλοί έχουν αναγνωρίσει ότι το στάδιο της προεργασίας δεδομένων είναι αυτό που επηρεάζει περισσότερο την επιτυχία ή όχι της ανεύρεσης γνώσης από δεδομένα .

Πρέπει επίσης, να σημειώσουμε ότι η διαδικασία των τεχνικών εξόρυξης από δεδομένα είναι επαναληπτική και συμμετοχική. Το αποτέλεσμα που προκύπτει από το κάθε στάδιο ή η επαναπληρόφρηση από τους ειδικούς που λαμβάνουν μέρος σε αυτή τη διαδικασία μπορεί να οδηγήσει σε διάφορες αλλαγές στα επόμενα στάδια.

Οι σημαντικότερες τεχνικές εξόρυξης από δεδομένα, οι οποίες αναφέρονται ως τύποι γνώσης που εξάγονται αυτόματα με τη μορφή κανόνων είναι:

- 1) Οι κανόνες συσχέτισης
- 2) Οι κανόνες ταξινόμησης
- 3) Οι κανόνες ομαδοποίησης

5.2. ΚΑΝΟΝΕΣ ΣΥΣΧΕΤΙΣΗΣ

5.2.1 ΓΕΝΙΚΑ

Οι κανόνες συσχέτισης (**association rules**) είναι μία από τις σημαντικότερες τεχνικές εξόρυξης από δεδομένα . Η κυριότερη εφαρμογή τους είναι η ανάλυση του καλαθιού της νοικοκυράς (**market-basket analysis**), όπου η αγορά ενός συνόλου προϊόντων (για παράδειγμα το περιεχόμενο ενός καλαθιού σούπερ-μάρκετ) αντιμετωπίζεται ως μια μεμονωμένη συναλλαγή.

Σκοπός είναι η εύρεση τάσεων (αλληλεξαρτήσεων / συσχετίσεων) μεταξύ ενός αριθμού συναλλαγών , οι οποίες στη συνέχεια μπορούν να αναλυθούν και μέσω αυτών να αναγνωριστούν αγοραστικά πρότυπα , τα οποία στη συνέχεια θα χρησιμοποιηθούν ως πληροφορία.

Η έννοια των κανόνων συσχέτισης εισήχθη το **1993** και αναφερόταν στην εξαγωγή συσχετίσεων στα πεδία βάσεων δεδομένων. Έκτοτε , έχει τύχει ευρείας ερευνάς και εφαρμογής . Η σημαντικότητα ενός κανόνα συσχέτισης καθορίζεται από το ποσοστό εφαρμογής του κανόνα στο σύνολο της βάσης δεδομένων (**Altman, E.L. 1993**) . Ένας κλασικός κανόνας συσχέτισης είναι για παράδειγμα της μορφής : Το **70%** των καταναλωτών που αγοράζουν ψωμί , αγοράζουν και βούτυρο , όπως επίσης και το **60%** των αγορών έχουν αγοραστεί ψωμί και βούτυρο).

Έτσι κάθε κανόνας συσχέτισης έχει δυο μεταβλητές μέτρησης , την επιβεβαίωση (**support**) και την αξιοπιστία (**confidence**). Αξιοπιστία μετράει ουσιαστικά την ισχύ του κανόνα, δηλαδή είναι ποσοστό των συναλλαγών που περιέχει το Y επί του αριθμού των συναλλαγών που περιέχουν το X. Η επιβεβαίωση αναφέρεται ουσιαστικά σε στατιστική σημαντικότητα και είναι το ποσοστό που εμφανίζονται το X και το Y μαζί

επί του συνόλου της βάσης δεδομένων (πόσο συχνά δηλαδή συμβαίνει το πρότυπο αυτό στη βάση δεδομένων). Το πρόβλημα εύρεσης κανόνων συσχέτισης αναφέρεται στην εύρεση εκείνων των κανόνων που έχουν μία καθορισμένη από τον χρήστη ελαχίστη επιβεβαίωση και αξιοπιστία.

Βασικά προβλήματα που αφορούν στην εύρεση κανόνων συσχέτισης αναφέρονται στην πολυπλοκότητα των αλγορίθμων και στη δυσκολία Επιλογής των χρήσιμων κανόνων που προκύπτουν. Το πρώτο πρόβλημα αναφέρεται στο γεγονός ότι ο αριθμός των κανόνων αυξάνεται εκθετικά με τον αριθμό των πεδίων. Οι σημερινοί αλγόριθμοι που εξάγουν κανόνες συσχέτισης μπορούν να μειώσουν αποτελεσματικά τον αριθμό αυτόν, βασιζόμενοι σε κατώφλια (**thresholds**), δηλαδή την επιβεβαίωση και την αξιοπιστία , για μέτρηση της ποιότητας των κανόνων. Το δεύτερο πρόβλημα αναφέρεται στο γεγονός ότι οι χρήσιμοι κανόνες που συνήθως προκύπτουν είναι μόνο ένα μικρό ποσοστό του συνόλου των κανόνων. Το πρόβλημα αυτό ερευνάται σε σχέση με την υποστήριξη προς τον χρήστη (**user support**), όταν αναζητά κανόνες μέσα στους κανόνες που έχουν εξαχθεί , καθώς και την ανάπτυξη επιπλέον μέτρων ποιότητας στους κανόνες .

Όπως ήδη αναφέρθηκε , ένα βασικό πρόβλημα στην εξαγωγή κανόνων συσχέτισης είναι ο τεράστιος αριθμός κανόνων που εξάγονται. Προκειμένου να μειωθεί ο αριθμός των κανόνων που εξάγονται, εισάγεται ο περιορισμός της ελάχιστης επιβεβαίωσης και της ελαχίστης επιβεβαίωσης και της ελάχιστης αξιοπιστίας (**misupp** και **minconf**) . Έτσι , το πρόβλημα εύρεσης κανόνων συσχέτισης μπορεί να διαιρεθεί στα δύο υπό-προβλήματα:

- 1) Εύρεση όλων των συνδυασμών των προϊόντων που έχουν επιβεβαίωση πάνω από την ελάχιστη επιβεβαίωση (η οποία καθορίζεται από το χρήστη). Όλοι αυτοί οι συνδυασμοί ονομάζονται μεγάλες λίστες από προϊόντα (**large itemsets**) και

όλοι οι υπόλοιποι συνδυασμοί μικρές λίστες από προϊόντα (**small itemsets**).

- 2) Εξόρυξη κανόνων συσχέτισης , βασισμένη στις μεγάλες λίστες από προϊόντα , οι οποίες έχουν μεγαλύτερη αξιοπιστία από αυτήν που καθορίζει ο χρήστης .

Έτσι, το πρόβλημα της εξόρυξης κανόνων συσχέτισης εντοπίζεται στην εύρεση των μεγάλων λιστών από προϊόντα δεδομένης μιας ελάχιστης επιβεβαίωσης.

5.2.2 ΑΛΓΟΡΙΘΜΟΙ ΣΥΣΧΕΤΙΣΗΣ

Οι αλγόριθμοι εξαγωγής κανόνων συσχέτισης μπορούν να χωρισθούν σε διάφορες κατηγορίες , ανάλογα με τον τύπο κανόνων που εξάγονται (κανονικούς-**Boolean**, χωρικούς , γενικευμένους , ποιοτικούς κ.α.), την αρχιτεκτονική του επεξεργαστή που χρησιμοποιείται (διαδοχικοί , παράλληλοι), τον τρόπο δημιουργίας υποψηφίων λιστών (για παράδειγμα, μια δυναμική στρατηγική δημιουργεί τις λίστες κατά τη διάρκεια του σκαναρίσματος των δεδομένων , ενώ μία υβριδική τεχνική δημιουργεί κάποιες υποψήφιες λίστες πριν και κάποιες κατά την διάρκεια), τη δομή των δεδομένων κ.α. Θα επικεντρωθούμε στους βασικότερους τύπους αλγορίθμων όπως αναλύονται παρακάτω.

5.2.3. ΔΙΑΔΟΧΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ (Sequential algorithms)

Στους διαδοχικούς αλγόριθμους γίνεται η υπόθεση ότι οι λίστες προϊόντων εντοπίζονται και αποθηκεύονται σε λεξιλογραφική σειρά , βασισμένη στο όνομα του προϊόντος. Αυτή η σειρά παρέχει ένα λογικό

τρόπο με τον οποίο τα προϊόντα και οι λίστες δημιουργούνται και μετρώνται. Οι διαδοχικοί αλγόριθμοι σχεδιάζονται να λειτουργούν με διαδοχικούς μηχανισμούς σε έναν κεντρικό επεξεργαστή. Ο αλγόριθμος αυτός στόχευε στην εξαγωγή ποιοτικών κανόνων. Στον αλγόριθμο αυτό οι υποψήφιας λίστες δημιουργούνται και μετρώνται καθώς σκανάρεται η βάση δεδομένων.

Μετά από την ανάγνωση μιας συναλλαγής, καθορίζεται ποια από τα προϊόντα που είχαν βρεθεί ως μεγάλες λίστες στην προηγούμενη επανάληψη περιλαμβάνονται στην συναλλαγή. Νέες υποψήφιας λίστες από προϊόντα δημιουργούνται με την επέκταση αυτών των μεγάλων λιστών με αλλά προϊόντα συναλλαγής.

Ο βασικός αλγόριθμος **Apriori** αναπτύχθηκε, προκειμένου να μειώσει τη δημιουργία τόσων πολλών υποψηφίων λιστών από προϊόντα (αυτό δηλαδή που έκαναν οι προηγούμενοι αλγόριθμοι), τα οποία τελικά μπορεί να ήταν μικρές λίστες. Η εύρεση των μεγάλων λιστών, βασίζεται στο ότι μια λίστα από προϊόντα είναι μεγάλη λίστα από προϊόντα, αν κάθε υποσύνολο της είναι μεγάλη λίστα από προϊόντα. Η εύρεση αυτή γίνεται κατόπιν πολλών επαναλήψεων στη βάση δεδομένων. Κατά την πρώτη επανάληψη, υπολογίζεται η επιβεβαίωση κάθε προϊόντος και υπολογίζεται ποια από αυτά είναι μεγάλες λίστες. Σε κάθε επόμενη επανάληψη λαμβάνονται υπόψη μόνο οι μεγάλες λίστες που είχαν βρεθεί στην προηγούμενη επανάληψη, χωρίς να λαμβάνονται υπόψη οι συναλλαγές. Από τις νέες λίστες δημιουργούνται νέες υποψήφιας μεγάλες λίστες.

Κατόπιν μετράται η επιβεβαίωση των λιστών αυτών και καθορίζεται ποιες από αυτές είναι τελικά μεγάλες λίστες.

Ο βασικός αλγόριθμος **Apriori** πλεονεκτεί, διότι μειώνει τη δημιουργία πολλών υποψηφίων λιστών από προϊόντα, είναι ο αλγόριθμος στον οποίο στηρίζονται οι περισσότεροι αλγόριθμοι που έχουν αναπτυχθεί έκτοτε και αποτελεί ένα μεγάλο βήμα στην ανάπτυξη των κανόνων συσχέτισης.

Μια παραλλαγή του βασικού αλγόριθμου **Apriori** είναι ο αλγόριθμος **Apriori TID**. Ένα ενδιαφέρον χαρακτηριστικό του αλγόριθμου αυτού είναι το γεγονός ότι δεν χρησιμοποιείται για υπολογισμό της επιβεβαίωσης των υποψηφίων μεγάλων λιστών η βάση δεδομένων **D** μετά την πρώτη επανάληψη. Αντί αυτού, χρησιμοποιείται μία κωδικοποίηση των υποψηφίων μεγάλων λιστών που είχε χρησιμοποιηθεί στην προηγούμενη επανάληψη. Σε επόμενες επαναλήψεις το μέγεθος της κωδικοποίησης αυτής μπορεί να γίνει πολύ μικρότερο από τον αριθμό των συναλλαγών στη βάση δεδομένων.

Μία επέκταση των **Apriori** και **AprioriTID** είναι ο αλγόριθμος **AprioriHybrid**, ο οποίος συνδυάζει τα καλύτερα χαρακτηριστικά και των δύο αλγόριθμων. Έτσι ο **AprioriHybrid** χρησιμοποιεί τον **Apriori** σε αρχικά στάδια και στη συνέχεια χρησιμοποιεί τον **AprioriTID**, όταν είναι πιθανό ότι η κωδικοποιημένη αρχική λίστα στο τέλος της επανάληψης θα χωράει στη μνήμη. Η αλλαγή αυτή από τον ένα αλγόριθμο στον άλλον, φυσικά περιλαμβάνει ένα κόστος.

Μια άλλη παραλλαγή του αλγόριθμου **Apriori** είναι ο αλγόριθμος **DIC (Dynamic itemset counting)**. Ο αλγόριθμος αυτός ομαλοποιεί τη διάκριση μεταξύ μετρήματος και δημιουργίας υποψηφίων λιστών.

Ο αλγόριθμος **PARTITION** χωρίζει τη βάση δεδομένων σε μικρά κομμάτια, έτσι ώστε να μπορούν να επεξεργάζονται ξεχωριστά και αποτελεσματικά, προκειμένου να βρεθούν οι μεγάλες λίστες.

Ο αλγόριθμος **OCD (Off-line Candidate Determination)** βασίζεται στην ιδέα ότι μικρά δείγματα συνήθως επιτυγχάνουν να βρουν μεγάλες λίστες με μεγαλύτερη ευκολία.

Ο αλγόριθμος **Sampling** μειώνει τις επαναλήψεις σε μία ή δύο. Ένα δείγμα αρχικά επιλέγεται από τη βάση δεδομένων, το οποίο μπορεί να χωρέσει στη βασική μνήμη. Στη συνέχεια οι μεγάλες λίστες βρίσκονται από δείγματα, χρησιμοποιώντας έναν βηματικό αλγόριθμο, όπως ο **Apriori**.

Ο αλγόριθμος **DHP (Direct Hashing and pruning)** είναι ιδιαίτερα αποτελεσματικός για την εύρεση υποψηφίων λιστών **2** προϊόντων.

Τέλος, ο αλγόριθμος **CARMA (Continuous Association Rule Mining Algorithm)**

Εμφανίζει **online** τους κανόνες συσχέτισης και επιτρέπει στο χρήστη να αλλάξει την ελάχιστη επιβεβαίωση και αξιοπιστία σε κάθε συναλλαγή κατά τη διάρκεια της πρώτης επανάληψης .

5.2.3.1 ΠΑΡΑΛΛΗΛΟΙ ΑΛΓΟΡΙΘΜΟΙ (Parallel algorithms)

Στην κατηγορία αυτή ανήκουν οι αλγόριθμοι που επικεντρώνονται στο πως θα παραλληλιστεί η λειτουργία της εύρεσης μεγάλων λιστών από προϊόντα και σχεδιάζονται για λειτουργία με παράλληλο τρόπο σε πολυεπεξεργαστές. Οι υπάρχοντες παράλληλοι αλγόριθμοι βασίζονται στο σειριακό αλγόριθμο **Apriori**.

Ο αλγόριθμος **CM (Count Distribution algorithm)** επικεντρώνεται στην ελαχιστοποίηση της επικοινωνίας και χρησιμοποιεί μία βασική αρχή επιτρέποντας « μειωμένους παράλληλους υπολογισμούς σε μη χρησιμοποιούμενος επεξεργαστές, προκειμένου να αποφευχθεί η επικοινωνία » Ο αλγόριθμος **DD (Data distribution algorithm)** επιχειρεί να αξιοποιήσει τη συνολική μνήμη του συστήματος πιο αποτελεσματικά. Τέλος, ο αλγόριθμος **Candidate Distribution** χωρίζει τόσο τα δεδομένα , όσο και τις υποψήφιες λίστες, με τέτοιο τρόπο , ώστε κάθε επεξεργαστής να μπορεί να δουλεύει ανεξάρτητα.

Ένας άλλος αλγόριθμος που ανήκει σε αυτή την κατηγορία είναι ο **PDM (Parallel Data Mining)**, ο οποίος προσπαθεί να παραλληλίσει τον **DHP (Park, C.,Han. I .2002)**.

Ο αλγόριθμος **DMA (Distributed Mining Algorithm)** εξάγει κανόνες από καταναμημένες βάσεις δεδομένων.

Ο αλγόριθμος **IDD (Intelligent Data Distribution)** αποτελεί μία βελτίωση του αλγόριθμου **DD**.

Επίσης ο αλγόριθμος **HD (Hybrid Distribution)** χωρίζει δυναμικά τις υποψήφιες λίστες, προκειμένου να διατηρήσει καλή ισορροπία κατά τη μεταφορά δεδομένων.

Τέλος ο αλγόριθμος **SH** δημιουργεί υποψήφιες λίστες ανεξάρτητα σε κάθε επεξεργαστή καθώς σκανάρει τη βάση δεδομένων και όχι εκ των προτέρων από τις προηγούμενες μεγάλες λίστες όπως ο σειριακός αλγόριθμος **Apriori**.

5.2.3.2 ΠΟΣΟΤΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ (Quantitative algorithms)

Οι αλγόριθμοι που αναπτύχτηκαν στα αρχικά στάδια έρευνας της εξαγωγής κανόνων συσχέτισης ασχολούνταν με κατηγορικά δεδομένα. Στην περίπτωση που τα δεδομένα είναι τόσο κατηγορικά, όσο και ποσοτικά, χρησιμοποιούνται οι ποσοτικοί αλγόριθμοι συσχέτισης, οι οποίοι χωρίζουν τις ποσότητες σε διαστήματα (**intervals**).

Εάν ένα ποσοτικό πρόβλημα μπορεί να μετατραπεί σε πρόβλημα με **Boolean** κανόνες, τότε κάθε αλγόριθμος για τη εύρεση κανονικών (**Boolean**) κανόνων συσχέτισης μπορεί να χρησιμοποιηθεί για την εύρεση ποσοτικών κανόνων συσχέτισης. Έτσι, οι τιμές των κατηγορικών μεταβλητών καταγράφονται από μια σειρά διαδοχικών ακέραιων αριθμών. Οι ποσοτικές μεταβλητές είτε χωρίζονται σε διαστήματα, όπου καταγράφονται σαν διαδοχικοί ακέραιοι, έτσι ώστε να τηρείται η σειρά των διαστημάτων, είτε δεν χωρίζονται σε διαστήματα, οπότε οι τιμές των μεταβλητών καταγράφονται σαν διαδοχικοί ακέραιοι, έτσι ώστε να διατηρείται η σειρά των τιμών.

Εάν τα διαστήματα είναι πολύ μικρά, μερικοί κανόνες μπορεί να μην έχουν ελάχιστη επιβεβαίωση. Από την άλλη μεριά, εάν τα διαστήματα είναι πολύ

μεγάλα, τότε μερικοί κανόνες μπορεί να μην έχουν ελάχιστη αξιοπιστία . Το πρόβλημα αυτό μπορεί να λυθεί με δύο τρόπους :

- a) Συνδυάζοντας κοντινά διαστήματα/τιμές .
- b) Εισάγοντας μια «μέγιστη επιβεβαίωση». Τα διαστήματα σταματούν να συνδυάζονται εάν η συνδυασμένη επιβεβαίωση τους ξεπερνάει την τιμή αυτή.

5.2.3.3 ΓΕΝΙΚΕΥΜΕΝΟΙ ΑΛΓΟΡΙΘΜΟΙ (Generalized algorithms)

Οι γενικευμένοι αλγόριθμοι χρησιμοποιούν δεδομένα με ιεραρχική δομή για την δημιουργία κανόνων συσχέτισης σε διαφορετικά επίπεδα της ιεραρχίας. Όταν οι κανόνες δημιουργούνται για προϊόντα σε υψηλότερο επίπεδο στην ιεραρχία, τόσο η επιβεβαίωση, όσο και η αξιοπιστία αυξάνονται. Το πρόβλημα της εύρεσης γενικευμένων κανόνων συσχέτισης αναφέρεται στη δημιουργία κανόνων για όλα τα επίπεδα. Επίσης, έχουν προταθεί διάφοροι παράλληλοι αλγόριθμοι για την δημιουργία γενικευμένων κανόνων συσχέτισης. Στιοπός είναι η δημιουργία μεγάλων λιστών με μέθοδο από πάνω προς τα κάτω (**top-down**) στην ιεραρχία, χρησιμοποιώντας έναν αλγόριθμο σαν τον **Apriori**.

5.2.3.4 ΧΩΡΙΚΟΙ ΚΑΙ ΧΡΟΝΟΛΟΓΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ (Spatial and temporal algorithms)

Οι χωρικοί αλγόριθμοι περιλαμβάνουν τοπικές πληροφορίες για την αποθήκευση των δεδομένων. Αυτό μπορεί να είναι ζευγάρια της μορφής γεωγραφικό μήκος- γεωγραφικό πλάτος , ταχυδρομικοί κώδικες, διευθύνσεις ή άλλα γεωγραφικά δεδομένα. Οι χωρικοί αλγόριθμοι συσχέτισης μπορούν να

φανούν ιδιαίτερα χρήσιμοι, όταν συνδυάζονται με Γεωγραφικά Συστήματα Πληροφοριών (ΓΣΠ).

Οι χρονολογικοί κανόνες συσχέτισης είναι παρόμοιοι με τους χωρικούς, με την διαφορά ότι συμπεριλαμβάνουν και τον παράγοντα χρόνο. Οι χωρικοί – χρονικοί κανόνες συμπεριλαμβάνουν τόσο χρόνο όσο και χώρο.

5.2.5 ΔΙΑΦΟΡΑ ΘΕΜΑΤΑ

Οι κανόνες συσχέτισης έχουν γενικευτεί σε κανόνες εξάρτησης οι οποίοι αναγνωρίζουν τη στατιστική εξάρτηση και σε περιπτώσεις απουσίας προϊόντων σε λίστες.

Επίσης έχει παρατηρηθεί ότι η πολυπλοκότητα στην εύρεση κανόνων συσχέτισης δεν εξαρτάτε μόνο από τον αριθμό των μεταβλητών, αλλά και από τον αριθμό των τιμών των μεταβλητών. Η πολυπλοκότητα αυτή μειώνεται με την ομαδοποίηση των τιμών σε διαστήματα..

Η μεγιστοποίηση στην εξαγωγή κανόνων συσχέτισης είναι ένα άλλο θέμα που έχει εμφανιστεί στην αντίστοιχη βιβλιογραφία. Η μεγιστοποίηση παρέχει έναν αποτελεσματικό τρόπο , προκειμένου να γίνει εστίαση στα πιο ενδιαφέροντα χαρακτηριστικά των μεταβλητών. Οι κανόνες συσχέτισης σε αυτή την περίπτωση μπορεί να περιλαμβάνουν απροσδιόριστες μεταβλητές. Το πρόβλημα που τίθεται είναι να προσδιοριστούν αυτές οι μεταβλητές, με τέτοιο τρόπο ώστε είτε η επιβεβαίωση είτε η αξιοπιστία να μεγιστοποιούνται **(Dimitras A.I. Slowinski.R.Susmaga. R. And Zopounidis. C. 1999).**

5.3 ΚΑΝΟΝΕΣ ΤΑΞΙΝΟΜΗΣΗΣ

5.3.1 ΓΕΝΙΚΑ

Η ταξινόμηση είναι ίσως η πιο διαδεδομένη τεχνική εξόρυξης από δεδομένα τα οποία είναι προταξινομημένα σε κλάσεις με σκοπό την ανάπτυξη ενός μοντέλου (πρότυπων) που θα μπορεί να ταξινομήσει νέα δεδομένα στις κλάσεις αυτές. Ένας αλγόριθμος ταξινόμησης χρησιμοποιεί αυτά τα προταξινομημένα δεδομένα προκειμένου να καθορίσει το σύνολο των παραμέτρων που χρειάζονται για περαιτέρω διάκριση (ταξινόμηση) δεδομένων. Ο αλγόριθμος στη συνέχεια κωδικοποιεί τις παραμέτρους (πρότυπα) αυτές σε ένα μοντέλο το οποίο ονομάζεται ταξινομητής (**Domingos. P. and Pazzani. M. 1997**)

Υπάρχουν διάφοροι τύποι αλγόριθμων ταξινόμησης: οι αλγόριθμοι που παράγουν δέντρα αποφάσεων (**decision trees**) (**Quinlan. J. R.1993**), αυτοί που βασίζονται σε νευρωνικά δίκτυα, σε λογιστική παλινδρόμηση, οι ταξινομητές **Bayes** και οι ταξινομητές **SVM (support vector machines)**. Τα μοντέλα **SVM** παράγουν διχοτομική ταξινόμηση το οποίο σημαίνει ότι δεν δίνεται η πιθανότητα βαθμού κλάσης (**probability of class membership**). Τα μοντέλα λογιστικής παλινδρόμησης (**logistic regression**) αποτελούνται από ανεξάρτητες μεταβλητές και από την εξαρτημένη μεταβλητή.

Τα νευρωνικά δίκτυα (**neural networks**) ακολουθούν την ίδια λογική με τη λογιστική παλινδρόμηση για την εκτίμηση των παραμέτρων, με τη διαφορά ότι τα όρια απόφασης (δηλαδή το όριο που καθορίζει τα σημεία στο χώρο που διαχωρίζουν τις δύο κλάσεις) μπορεί να είναι μη γραμμικά. Επίσης, οι ταξινομητές **Bayes** θεωρούν ότι τα χαρακτηριστήρια είναι ανεξάρτητα μεταξύ τους (δεδομένης της κλάσης), και αποθηκεύουν μία πιθανολογική περίληψη για κάθε κλάση, προκειμένου να κάνουν την ταξινόμηση.

Όσον αφορά τους κανόνες και τα δέντρα αποφάσεων, και οι δύο μέθοδοι μοιάζουν πολύ μεταξύ τους, με τη διαφορά ότι τα δέντρα αποφάσεων

παράγουν και μία οπτική παρουσίαση των κανόνων, κάτι το οποίο τα κάνει πιο δεδομένα ως μέθοδο για ταξινόμηση.

Ένα δέντρο αποφάσεων κατασκευάζεται από ένα σύνολο εκπαίδευσης (**training set**), δηλαδή από ένα σύνολο δεδομένων / εγγραφών. Κάθε εγγραφή χαρακτηρίζεται από ένα σύνολο χαρακτηριστικών (**attributes**) και μία κλάση (**label**). Η λογική της κατασκευής ενός δέντρου αποφάσεων είναι η σωστή και ακριβής σχέση(ή αλληλεξάρτηση)των χαρακτηριστικών αυτών και της κλάσης. Ένα δέντρο αποφάσεων περιέχει μηδενικούς ή περισσότερους ενδιάμεσους κόμβους (**internal nodes**) και έναν ή περισσότερους τερματικούς (**leaf**) κόμβους. Κάθε ενδιάμεσος κόμβος αποτελείται από δύο ή περισσότερους κόμβους – παιδιά (**child nodes**). Όλοι οι ενδιάμεσοι κόμβοι περιέχουν διαιρέσεις (**splits**), οι οποίες ελέγχουν την τιμή της έκφρασης των χαρακτηριστικών. Τέλος, ένας τερματικός κόμβος αποτελείται από μία τιμή κλάσης.

Οι βασικοί αντικειμενικοί σκοποί των ταξινομητών δέντρων αποφάσεων είναι:

1. Να ταξινομήσουν σωστά όσο το δυνατόν περισσότερο ποσοστό από το σύνολο εκπαίδευσης (**training set**), δηλαδή τα νέα δεδομένα .
2. Να γενικεύσουν πέρα από το δείγμα εκπαίδευσης, έτσι ώστε ένα νέο και άγνωστο δείγμα εκπαίδευσης, να μπορεί να ταξινομηθεί με όσο το δυνατό μεγαλύτερη ακρίβεια.
3. Να μπορούν να κάνουν ενημέρωση (**update**), όταν γίνονται διαθέσιμα περισσότερα δεδομένα.
4. Να έχουν όσο πιο απλή δομή γίνεται.

5.3.2 ΑΛΓΟΡΙΘΜΟΙ ΤΑΞΙΝΟΜΗΣΗΣ ΔΕΝΤΡΩΝ

Οι αλγόριθμοι που παράγουν δέντρα αποφάσεων ακολουθούν συνήθως αναλυτική προσέγγιση, δηλαδή παραγωγή (**induction**) δέντρων από πάνω προς τα κάτω (**top-down**).

Δημιουργούν δηλαδή το δέντρο από τη ρίζα και συνεχίζουν προς τα κάτω, επιλέγοντας ένα πεδίο ή χαρακτηριστικό (**attribute**) από όλο το σύνολο των χαρακτηριστικών, το οποίο και τοποθετείται στη ρίζα του δέντρου. Στη συνέχεια για κάθε τιμή (ή διάστημα) του χαρακτηριστικού αυτού ορίζεται ένα υποσύνολο εγγραφών, οι εγγραφές του οποίου έχουν το συγκεκριμένο χαρακτηριστικό τη συγκεκριμένη τιμή (ή διάστημα). Αφού ολοκληρωθεί το βήμα αυτό και ο αλγόριθμος έχει κάνει την πρώτη διακλάδωση, αναζητά για κάθε υποσύνολο ένα υποδέντρο αποφάσεων (**sudtree**). Όταν βρει ένα υποσύνολο, το οποίο ανήκει αποκλειστικά σε μία μονή κλάση, τότε η διαδικασία σταματά, η διακλάδωση προς τα κάτω τελειώνει και παίρνει φύλλο με την κλάση, στην οποία ανήκει το υποσύνολο.

Αξίζει να σημειωθεί ότι έχουν προταθεί και άλλες προσεγγίσεις για το σχεδιασμό ταξινομητών δέντρων αποφάσεων, όπως η προσέγγιση **bottom-up**, όπου υπολογίζονται οι αποστάσεις μεταξύ προταξινομημένων κλάσεων και σε κάθε βήμα οι δύο κλάσεις με την μικρότερη απόσταση ενώνονται, ώστε να δημιουργήσουν μια νέα ομάδα, μέχρις ότου μείνει ένας κόμβος ο οποίος περιέχει όλες τις κλάσεις, δηλαδή η ρίζα του δέντρου. Επίσης, έχει προταθεί μια υβριδική (**hybrid**) μέθοδος η οποία συνδυάζει τόσο αναλυτική (**top-down**) προσέγγιση, όσο και προσέγγιση **bottom-up** διαδοχικά. Παρ' όλα αυτά η πιο διαδεδομένη μέθοδος σχεδιασμού ταξινομητών δέντρων αποφάσεων είναι η αναλυτική προσέγγιση, η οποία αναλύεται στα παρακάτω στάδια:

1. Η επιλογή κριτηρίου επιλογής του πεδίου για κάθε κόμβο, βάσει του οποίου θα γίνει η διακλάδωση (**splitting rules**)
2. Η απόφαση σχετικά με το ποιοι κόμβοι είναι τερματικοί.
3. Η αντιστοιχία κάθε τερματικού κόμβου με ένα φύλλο κλάσης.

Μια εγγραφή ταξινομείται, αφού περάσει από το δέντρο ξεκινώντας από τη ρίζα. Ο έλεγχος σε κάθε ενδιάμεσο κόμβο εφαρμόζεται στα χαρακτηριστικά της εγγραφής προκειμένου να καθοριστεί το επόμενο

τόξο (**arc**), στο οποίο η εγγραφή πρέπει να προχωρήσει. Η τιμή στον τερματικό κόμβο, στον οποίο καταλήγει η εγγραφή είναι και η ταξινόμηση της. Μία εγγραφή. Το ποσοστό των εγγραφών που ταξινομείται λάθος (**misclassified**) από το δέντρο, εάν η ταξινόμηση της δεν είναι η ίδια από τη σωστή κλάση της εγγραφής. Το ποσοστό των εγγραφών που ταξινομείται σωστά από ένα δέντρο αποφάσεων ονομάζεται ακρίβεια (**accuracy**), ενώ το ποσοστό των λανθασμένων ταξινομημένων εγγράφων αναφέρεται ως λάθος (**error**).

Στη συνέχεια ελέγχεται η ακρίβεια της ταξινόμησης. Έτσι, αν όλες οι εγγραφές έχουν ταξινομηθεί σωστά, ο αλγόριθμος τερματίζει, διαφορετικά προστίθενται και άλλες εγγραφές στο παράθυρο και η διαδικασία επαναλαμβάνεται, μέχρις ότου όλες οι εγγραφές ταξινομηθούν σωστά από το δέντρο. Βασική παράμετρος του αλγόριθμου είναι ποιο ποσοστό των εγγραφών θα περιέχει το παράθυρο και με ποιο ρυθμό θα μεγαλώνει, εφόσον δεν είναι επαρκές.

Σημαντικότερη παράμετρο στον αλγόριθμο αποτελεί το κριτήριο επιλογής του πεδίου για κάθε κόμβο, βάση του οποίου θα γίνει η διακλάδωση. Ο αλγόριθμος αυτός χρησιμοποιεί σαν κριτήριο επιλογής την Εντροπία, η οποία δίνει μία εκτίμηση του πόσο λανθασμένα χωρίζεται κάθε φορά το σύνολο εκπαίδευσης, βάση του συγκεκριμένου πεδίου. Η εντροπία είναι ένα μέγεθος που χρησιμοποιείται στη Θεωρία της Πληροφορίας .

5.3.3 ΚΑΝΟΝΕΣ ΔΙΑΚΛΑΔΩΣΗΣ / ΔΙΑΙΡΕΣΗΣ ΔΕΝΤΡΩΝ (Splitting rules)

Όπως αναφέρθηκε σε προηγούμενη παράγραφο, ένα από τα βασικότερα στοιχεία των ταξινομητών δέντρων αποφάσεων είναι η εύρεση ενός τεστ- κανόνα που θα διαιρεί τα δεδομένα σε υποομάδες σε

κάθε ενδιάμεσο κόμβο. Έτσι βασικό ρόλο στην παραγωγή δέντρων αποφάσεων έχει η κατάταξη των χαρακτηριστικών βάσει της χρησιμότητας τους στο διαχωρισμό κλάσεων στα δεδομένα..

Σε περιπτώσεις δέντρων αποφάσεων που χρησιμοποιούν κανόνες διακλάδωσης βασισμένους σε ένα μόνο χαρακτηριστικό σε κάθε ενδιάμεσο κόμβο (τα δέντρα αυτά ονομάζονται **univariate** και αποτελούν την πιο συνηθισμένη περίπτωση), χρησιμοποιούνται διάφοροι κανόνες εκτίμησης των χαρακτηριστικών (**feature evaluation rules**), δανεισμένοι είτε από τη Θεωρία της Πληροφορίας (**information theory**) είτε βασισμένοι σε μετρικές αποστάσεις (**distance measures**), είτε σε μεγέθη εξάρτησης (**dependence measures**).

Στην πρώτη περίπτωση ανήκει η εντροπία, όπως αναφέρθηκε σε προηγούμενη παράγραφο. Η εντροπία έχει εφαρμοστεί στην αναγνώριση προτύπων για την κατασκευή δέντρων, με σκοπό την μεγιστοποίηση της συνολικής αμοιβαίας πληροφορίας, δηλαδή με την επέκταση κόμβων που συνεισφέροντας μεγαλύτερο όφελος στη μέση αμοιβαία πληροφόρηση σε όλο το δέντρο.

Στη δεύτερη κατηγορία, η απόσταση αναφέρεται ως η απόσταση μεταξύ κλάσεων πιθανολογικών κατανομών (**class probability distribution**). Σαν κανόνες διαίρεσης έχουν προταθεί ο δείκτης **Gini** διαφοροποίησης (**Gini index of diversity**), και περιλαμβάνεται στον αλγόριθμο **CART**, ο οποίος έχει χρησιμοποιηθεί για κατασκευή δέντρων στη στατιστική, καθώς και στην αναγνώριση προτύπων. Στην τρίτη κατηγορία ανήκουν οι κανόνες οι οποίοι μετρούν τη στατιστική εξάρτηση μεταξύ δυο τυχαίων μεταβλητών. Για την κατασκευή δέντρων αποφάσεων έχει χρησιμοποιηθεί ένας συνδυασμός του στατιστικού X^2 και της αμοιβαίας πληροφόρησης (**mutual information**). Στην περίπτωση αυτή, μετράται το όφελος στη μέση αμοιβαία πληροφόρηση εξαιτίας μιας καινούριας διαίρεσης και κατόπιν υπολογίζεται η

πιθανότητα το όφελος αυτό να προέρχεται από τύχη, χρησιμοποιώντας πίνακες κατανομής X^2 . Έτσι επιλέγεται σα διαίρεση εκείνη, η οποία ελαχιστοποιεί αυτή την πιθανότητα. Ως κανόνες διαίρεσης έχουν επίσης προταθεί ο συνολικός αριθμός των λανθασμένα ταξινομημένων εγγραφών, όπως η αθροιστική μειοψηφία (**sum minority**) και η ακρίβεια (**inaccuracy**), καθώς επίσης και ο μέγιστος αριθμός των λανθασμένα ταξινομημένων εγγραφών σε δύο πλευρές μιας διαίρεσης με δύο κλάσεις.

Οι περισσότεροι από τους παραπάνω κανόνες διαίρεσης υποθέτουν ότι δεν είναι γνωστή η πιθανολογική κατανομή των εγγράφων. Σε αντίθεση με αυτούς, έχουν προταθεί κανόνες, όπου υποθέτεται πλήρης γνώση των πιθανοτήτων των δεδομένων, όπως βέλτιστοι κανόνες σε κάθε κόμβο του δέντρου (κανόνες που ελαχιστοποιούν το συνολικό πιθανολογικό λάθος). Έτσι δέντρα που παράγονται άμεσα από τιμές των χαρακτηριστικών.

Στις περιπτώσεις δέντρων αποφάσεων που χρησιμοποιούν κανόνες διακλάδωσης λαμβάνοντας υπόψη περισσότερα από ένα χαρακτηριστικά σε κάθε ενδιάμεσο κόμβο (τα δέντρα αυτά ονομάζονται **multivariate**), χρησιμοποιούνται (παράγονται) γραμμικά δέντρα, δηλαδή δέντρα, των οποίων τα τεστ στους ενδιάμεσους κόμβους βασίζονται σε ένα γραμμικό συνδυασμό των χαρακτηριστικών. Το πρόβλημα εύρεσης βέλτιστων διακλαδώσεων στην περίπτωση αυτή είναι πιο δύσκολο από την περίπτωση των **univariate** δέντρων, έτσι είναι απαραίτητες κάποιες ευρετικές μέθοδοι (**heuristics**), προκειμένου να βρεθούν οι βέλτιστες γραμμικές διακλαδώσεις.

Ο γραμμικός προγραμματισμός είναι μία άλλη μέθοδος, η οποία έχει κατά καιρούς χρησιμοποιηθεί για εύρεση διακλαδώσεων στην ταξινόμηση. Σκοπός της μεθοδολογίας αυτής είναι η ελαχιστοποίηση της αποστάσεις των λανθασμένα ταξινομημένων σημείων από το όριο

της απόφασης (**decision boundary**). Έτσι δεδομένων δύο πιθανών ομάδων σημείων, έχει προταθεί ένα μοντέλο γραμμικού προγραμματισμού για εύρεση εκείνης της διακλάδωσης, της οποίας η απόσταση από τα λανθασμένα ταξινομημένα σημεία ελαχιστοποιείται.

Έχει επίσης χρησιμοποιηθεί ακέραιος προγραμματισμός για τον ίδιο σκοπό. Τα νευρωτικά δίκτυα έχουν επίσης χρησιμοποιηθεί σε συνδυασμό με τα δέντρα αποφάσεων. Το αποτέλεσμα που προκύπτει μπορεί να ερμηνευτεί ως δέντρα αποφάσεων με μη γραμμικές διακλαδώσεις. Μία μεθοδολογία μετατρέπει ένα **univariate** δέντρο σε νευρωτικό δίκτυο, στη συνέχεια εκπαιδεύεται και καταλήγει σε δίκτυα εντροπίας με δομή δέντρων.

Επίσης έχουν χρησιμοποιηθεί δέντρα αποφάσεων με μικρά πολλαπλά δίκτυα σε κάθε κόμβο, τα οποία παράγουν μη γραμμικές, πολλαπλές διακλαδώσεις.

Τέλος αξίζει να σημειωθεί ότι οι αρχικοί αλγόριθμοι ταξινόμησης έχουν επεκταθεί, (οι οποίοι αρχικά είχαν σχεδιαστεί για συνεχείς αριθμητικές μεταβλητές), ώστε να συμπεριλαμβάνουν και διακριτές τιμές αριθμητικών μεταβλητών, καθώς και κατηγορικές μεταβλητές με πολλαπλές τιμές.

5.3.4 ΜΕΘΟΔΟΙ ΠΕΡΙΟΡΙΣΜΟΥ ΕΚΤΑΣΗΣ ΤΩΝ ΔΕΝΤΡΩΝ

Ένα βασικό πρόβλημα στην παραγωγή δέντρων αποφάσεων που έχει απασχολήσει τους ερευνητές είναι η επιλογή του σημείου, στο οποίο το δέντρο σταματά να επεκτείνεται ή αλλιώς το κατάλληλο μέγεθος του δέντρου.

Έχει αναφερθεί ότι η ποιότητα ενός δέντρου εξαρτάται περισσότερο από κανόνες που ορίζουν σωστά που το δέντρο σταματά, παρά από τους

κανόνες διακλάδωσης. Η πιο γνωστή μέθοδος για την παραγωγή κατάλληλου μεγέθους ενός δέντρου απόφασης είναι η μέθοδος περικοπής (Papoulias, C., and Theodossiou, P.,1992).

Η μέθοδος αυτή λειτουργεί ως εξής : αρχικά κατασκευάζεται ολόκληρο το δέντρο, δηλαδή το δέντρο εκείνο , στο οποίο καμία περαιτέρω διακλάδωση σε τερματικό κόμβο δε βελτιώνει την ακρίβεια (**accuracy**) στο σύνολο εκπαίδευσης. Στη συνέχεια, καταργούνται τα υπόδεντρα τα οποία δεν συνεισφέρουν σημαντικά στην γενικευμένη ακρίβεια (**generalization accuracy**).

Έχει διατυπωθεί ότι το ξεχωριστό σύνολο για περικοπή μπορεί να δημιουργήσει προβλήματα, ιδιαίτερα όταν το σύνολο εκπαίδευσης είναι μικρό. Για αυτό το λόγο, έχουν προταθεί διάφορες μεθοδολογίες, όπως αυτή της διαδικασίας **cross validation**, η οποία δεν κρατάει ένα υποσύνολο των δεδομένων για περικοπή, αλλά παρ' όλα αυτά εμπεριέχει μεγάλη αριθμητική πολυπλοκότητα. Εναλλακτικά, προτείνεται μια διαδοχική μέθοδος, η οποία χωρίζει το σύνολο εκπαίδευσης σε δύο μέρη, εκ των οποίων το ένα χρησιμοποιείται για την παραγωγή του δέντρου και το άλλο για περικοπή με εναλλασσόμενους ρόλους σε κάθε επανάληψη.

Εκτός από τη μέθοδο περικοπής, όπως αναφέρθηκε, έχουν εφαρμοσθεί και άλλες τεχνικές για την παραγωγή κατάλληλου μεγέθους δέντρων αποφάσεων. Αυτές αναφέρονται ως ακολούθως :

- ✓ Περιορισμοί στο ελάχιστο μέγεθος των κόμβων : Έχει χρησιμοποιηθεί σε αρχικές μεθοδολογίες. Ένας κόμβος δεν διακλαδώνεται, εάν έχει λιγότερες από **k** εγγραφές, όπου το **k** αποτελεί παράμετρο στον αλγόριθμο.
- ✓ Μέθοδος δύο σταδίων: Αρχικά καθορίζεται μια καλή δομή για το δέντρο. Κατόπιν, βρίσκονται διακλαδώσεις σε όλους τους

κόμβους. Για το πρώτο στάδιο έχει χρησιμοποιηθεί πολλαπλή διαδοχική

- ▼ παλινδρόμηση , ενώ για το δεύτερο ανάλυση γραμμικής διακριτής ταξινόμησης.

5.3.5 ΣΥΓΚΡΙΣΗ ΜΕ ΑΛΛΕΣ ΜΕΘΟΔΟΛΟΓΙΕΣ

Τα δέντρα αποφάσεων έχουν κατά καιρούς συγκριθεί με άλλες ανταγωνιστικές μεθόδους ανάλυσης δεδομένων και ταξινόμησης στη στατιστική και στην επιστήμη γνώσεων από μηχανές (**machine learning**), προκειμένου να μετρηθεί η αποτελεσματικότητά τους έναντι αυτών των μεθόδων. Σε γενικές γραμμές έχει αναφερθεί ότι πολλές από τις μεθόδους αυτές παράγουν μεν ακριβείς ταξινομητές, αλλά δεν παρέχουν αρκετή πληροφόρηση σχετικά με τη δομή του προβλήματος.

Έτσι τα δέντρα αποφάσεων έχουν συγκριθεί με νευρωνικά δίκτυα, όπου τα αποτελέσματα δείχνουν ότι δεν υπάρχει μεγάλη διαφορά μεταξύ των δύο μεθοδολογιών ως προς την ποιότητα των αποτελεσμάτων. Σε μια άλλη μελέτη τα δέντρα αποφάσεων συγκρίνονται με νευρωνικά δίκτυα, όπου παρουσιάζεται ότι τα δέντρα αποφάσεων είναι πολύ γρηγορότερα από τα νευρωνικά δίκτυα για την εκτέλεση συγκεκριμένων λειτουργιών. Παρ' όλα αυτά έχει επίσης διατυπωθεί ότι τα δέντρα αποφάσεων δεν μπορούν να λάβουν υπόψη τους κάποιες στατιστικές πληροφορίες, οι οποίες είναι διαθέσιμες στα νευρωνικά δίκτυα.

Γενικότερα, έχει αναφερθεί ότι τα βασικά πλεονεκτήματα των ταξινομητών δέντρων είναι τα ακόλουθα :

1. Η γνώση που αποκτάται από προταξινομημένα παραδείγματα ξεπερνάει το "μποτιλιάρισμα" της απόκτησης γνώσης από έναν **domain expert**.

2. Οι μέθοδοι εξαγωγής δέντρων αποφάσεων είναι διερευνητικές και όχι επαγωγικές. Επίσης είναι μη- παραμετρικές, αρκούν μερικές παραδοχές σχετικά με το μοντέλο και την κατανομή των δεδομένων, ώστε τα δέντρα να μοντελοποιήσουν ένα μεγάλο εύρος κατανομής δεδομένων.
3. Η ιεραρχική διάκριση που παράγουν τα δέντρα αποφάσεων προσφέρει καλύτερη και αποτελεσματικότερη χρήση των διαθέσιμων χαρακτηριστικών (**features**).
4. Τα δέντρα αποφάσεων μπορούν να χρησιμοποιηθούν τόσο σε προβλήματα με διακριτές τιμές, όσο και σε ημιτελή προβλήματα (σε διακριτά προβλήματα η εξαρτημένη μεταβλητή μπορεί να προβλεφτεί απόλυτα από τις ανεξάρτητες μεταβλητές, κάτι το οποίο δεν ισχύει σε ημιτελή προβλήματα).
5. Τα δέντρα αποφάσεων ταξινομούν τα δεδομένα με βάση μια σειρά εύκολων και κατανοητών κριτηρίων.
6. Σε σχέση με κάποιες στατιστικές μεθόδους, οι ταξινομητές δέντρων αποφάσεων χειρίζονται τα **unimodal** και **multimodal** δεδομένα με τον ίδιο τρόπο.

Οι ταξινομητές δέντρων αποφάσεων έχουν επίσης διάφορα μειονεκτήματα, τα οποία συνοψίζονται ως εξής :

1. Κάποια λάθη μπορεί να συγκεντρώνονται από επίπεδο σε επίπεδο ένα μεγάλο δέντρο. Είναι λοιπόν εξαιρετικά δύσκολο να μεγιστοποιηθεί τόσο η ακρίβεια (**accuracy**) όσο και η αποτελεσματικότητα (**efficiency**). Έτσι, για κάποια δεδομένη ακρίβεια, πρέπει να ικανοποιείται ένα ποσοστό αποτελεσματικότητας.
2. Οι επικαλύψεις (**overlap**) (όταν δηλαδή δύο ενδιάμεσοι κόμβοι έχουν μια κοινή κλάση) μπορεί να προκαλέσουν μεγάλο αριθμό των κλάσεων, κάτι το οποίο αυξάνει το χρόνο κατασκευής και το χώρο μνήμης που χρειάζεται για την επεξεργασία.

3. Επίσης, μπορεί να υπάρχουν δυσκολίες σχετικά με το σχεδιασμό ενός ταξινομητή δέντρου αποφάσεων (Platt, J., 1999).

5.4 ΚΑΝΟΝΕΣ ΟΜΑΔΟΠΟΙΗΣΗΣ

5.4.1 ΓΕΝΙΚΑ

Οι κανόνες ομαδοποίησης αναφέρονται στο διαχωρισμό (ομαδοποίηση) δεδομένων σε ομογενείς ομάδες (**clusters**), έτσι ώστε οι εγγραφές που βρίσκονται σε κάθε ομάδα να είναι πιο όμοιες από ό,τι άλλες εγγραφές που βρίσκονται σε άλλες ομάδες, με βάση ορισμένα προκαθορισμένα κριτήρια. Σκοπός των τεχνικών αυτών είναι η μεγιστοποίηση της συνοχής και ομοιότητας μέσα στις ομάδες βάση των χαρακτηριστικών (**attributes**) των εγγραφών τους και η ελαχιστοποίηση της ομοιότητας ανάμεσα στις εγγραφές των διαφορετικών ομάδων. Η βασική διαφορά μεταξύ των κανόνων ταξινόμησης είναι ότι τα δεδομένα που χρησιμοποιούνται για μάθηση στους κανόνες ομαδοποίησης δεν είναι προταξινομημένα (**supervised**), δηλαδή δεν υπάρχει κάποια ιδιότητα – στόχος, ενώ τα δεδομένα που χρησιμοποιούνται για μάθηση στους κανόνες ταξινόμησης είναι προταξινομημένα βάσει της ιδιότητας (κλάσης) αυτής. Έτσι, στους κανόνες ομαδοποίησης δεν υπάρχουν προταξινομημένες κλάσεις που να υποδεικνύουν τις επιθυμητές σχέσεις ανάμεσα στα δεδομένα. Μερικές από τις τεχνικές εξόρυξης από δεδομένα που χρησιμοποιούν τεχνικές ομαδοποίησης είναι ο μερισμός (**segmentation**) βάσεων δεδομένων, τα μοντέλα πρόβλεψης και η αστικοποίηση δεδομένων (**visualization**) σε μεγάλες βάσεις δεδομένων.

Η διαδικασία των τεχνικών ομαδοποίησης συνοψίζεται στα ακόλουθα στάδια:

- 1) Επιλογή κριτηρίων (**feature selection**):

Σκοπός είναι η επιλογή κριτηρίων, δηλαδή η αναγνώριση των πιο αποτελεσματικών υποσύνολο κριτηρίων, βάσει των οποίων θα εφαρμοσθεί η ομαδοποίηση στα δεδομένα, έτσι ώστε να συμπεριληφθούν όσο το δυνατόν περισσότερες πληροφορίες.

2) Αλγόριθμος ομαδοποίησης (**clustering algorithm**):

Σε αυτό το στάδιο επιλέγεται ο κατάλληλος αλγόριθμος που θα χωρίσει τα δεδομένα σε ομάδες. Ένα μέτρο γειννίας (**proximity measure**), καθώς και ένα κριτήριο ομαδοποίησης (**clustering criterion**) χαρακτηρίζουν έναν αλγόριθμο ομαδοποίησης, καθώς και η αποτελεσματικότητά του στο να ορίζει ένα σχήμα (**scheme**) ομαδοποίησης που να ταιριάζει στα δεδομένα.

§ Μέτρο γειννίας/ εγγύτητας (**proximity measure**): Το μέτρο γειννίας είναι ένα μέτρο που πολιτικοποιεί πόσο όμοια είναι δύο σημεία δεδομένων (εγγραφές) και συνήθως μετράται από μία συνάρτηση απόστασης (**distance function**) που εφαρμόζεται σε αυτά τα σημεία.

§ Κριτήριο ομαδοποίησης (**clustering criterion**) : Το στάδιο αυτό αναφέρεται στην επιλογή του κριτηρίου ομαδοποίησης, το οποίο εκφράζεται μέσω μιας συνάρτησης κόστους ή αλλιώς.

3) Τεκμηρίωση των αποτελεσμάτων (**validation of results**):

Τα αποτελέσματα που προκύπτουν από τη διαδικασία της ομαδοποίησης πρέπει να τεκμηριωθούν με διάφορες τεχνικές ή κριτήρια. Όμως, επειδή οι κανόνες ομαδοποίησης παράγουν ομάδες – κανόνες που δεν είναι γνωστές εκ των προτέρων (**a priori**), δεν υπάρχουν κοινώς αποδεκτά πρότυπα τεκμηρίωσης. Κατά συνέπεια, ο μόνος τρόπος να τεκμηριωθούν τα αποτελέσματα είναι μέσω της διαδικασίας τεκμηρίωσης, η οποία εφαρμόζεται στα αποτελέσματα με σκοπό να φανεί αν αυτά έχουν αξία και νόημα ή όχι.

Έτσι μπορούν να χρησιμοποιηθούν στατιστικές μέθοδοι, ώστε να ελέγχουν στατιστικές υποθέσεις, εξωτερική αξιολόγηση, η οποία συγκρίνει τη δομή που

έχει προκύψει με μία προκαθορισμένη (**a priori**), δομή, ή και εσωτερική αξιολόγηση, η οποία αξιολογεί πόσο η προκύπτουσα δομή είναι κατάλληλη για τα δεδομένα.

4) Επεξήγηση των αποτελεσμάτων (**interpretation of the results**):

Σε μερικές περιπτώσεις οι χρήστες πρέπει να επεξηγήσουν τα αποτελέσματα που προκύπτουν, λαμβάνοντας υπόψη και άλλα αποτελέσματα από άλλες διαδικασίες, δηλαδή να ενώσουν τα αποτελέσματα που προκύπτουν από την ομαδοποίηση με άλλα πειραματικά αποτελέσματα, ώστε να οδηγηθούν στο σωστό συμπέρασμα.

Γενικά, οι τεχνικές ομαδοποίησης αποσκοπούν στους εξής σκοπούς :

1. Συμπίεση των δεδομένων (**data reduction**): Οι τεχνικές ομαδοποίησης μπορούν να προσδιορίσουν ‘ενδιαφέρουσες’ ομάδες σε βάσεις δεδομένων σε περιπτώσεις όπου η ανάλυση των δεδομένων αυτών είναι πολύ δύσκολη λόγω του τεράστιου όγκου τους.
2. Δημιουργία υποθέσεων : Οι τεχνικές ομαδοποίησης μπορούν να αναγνωρίσουν υποθέσεις σχετικά με τα δεδομένα. Κατά την διαδικασία ομαδοποίησης μπορεί να εξαχθεί κάποιος κανόνας (για παράδειγμα, ομαδοποίησης πελατών ανάλογα με την ηλικία και τον χρόνο αγοράς) που επιτρέπει τη διατύπωση μιας υπόθεσης (οι νέοι πηγαίνουν για ψώνια το απόγευμα, άρα οι μεγαλύτεροι σε ηλικία πηγαίνουν για ψώνια το πρωί).
3. Έλεγχος υποθέσεων : Σε αυτή την περίπτωση, οι τεχνικές ομαδοποίησης ελέγχουν και αξιολογούν κάποιες υποθέσεις σε σχέση με τα δεδομένα. Έτσι μπορεί να δημιουργηθεί μια γενική διατύπωση για ένα συγκεκριμένο κομμάτι της αγοράς.
4. Ανεύρεση γνώσης και πρόβλεψης βάση των δεδομένων που έχουν διαμορφωθεί: Η ομαδοποίηση δεδομένων μπορεί να δώσει

πληροφορίες σχετικά με τα χαρακτηριστικά των δεδομένων που βρίσκονται σε αυτές.

Οι κανόνες ομαδοποίησης διακρίνονται σε διάφορες κατηγορίες ανάλογα με τα έξι κριτήρια :

- § Ο τύπος των εισερχόμενων δεδομένων στον αλγόριθμο.
- § Το κριτήριο ομαδοποίησης, το οποίο ορίζει την ανομοιότητα ή ομοιότητα ανάμεσα σε δύο εγγραφές.
- § Θεωρία και οι βασικές αρχές, στις οποίες στηρίζονται οι τεχνικές ομαδοποίησης, όπως για παράδειγμα η στατιστική και η θεωρία **fuzzy**.

Έτσι, ανάλογα με τη μέθοδο που χρησιμοποιείται για να ορισθούν οι ομάδες, οι αλγόριθμοι ομαδοποιήσεως μπορούν να ταξινομηθούν στις ακόλουθες κατηγορίες :

- § Διαχωριστική ομαδοποίηση (**partitional clustering**)
- § Ιεραρχική ομαδοποίηση (**hierarchical clustering**)
- § Ομαδοποίηση βάσει συνθηκών πυκνότητας (**Density-based**)
- § Ομαδοποίηση βάσει πλεγμάτων (**Grid-based**)

Επίπλυν. Για κάθε μία από τις παραπάνω κατηγορίες υπάρχουν διάφορες υποκατηγορίες και διάφοροι αλγόριθμοι. Έτσι, η τεχνική της ομαδοποίησης μπορεί να διαχωριστεί ανάλογα με τον τύπο των μεταβλητών στα δεδομένα ως εξής :

- § Στατιστική ή αριθμητική (**numerical**) : Στην περίπτωση αυτή χρησιμοποιούνται αριθμητικά κριτήρια ομοιότητας. Οι ομάδες που προκύπτουν περιγράφονται από αριθμητικές τιμές.
- § Εννοιολογική (**conceptual**) : Ο προσδιορισμός των ομάδων βασίζεται στο νόημα και στις έννοιες

που αντιπροσωπεύουν και χρησιμοποιείται για κατηγορικά δεδομένα.

Τέλος, ανάλογα με τον τροποποιώ χειρίζονται την αβεβαιότητα, οι τεχνικές ομαδοποίησης διακρίνονται σε **fuzzy** ομαδοποίηση, όπου κάθε εγγραφή κατηγοριοποιείται σε περισσότερες από μία ομάδες, ανάλογα με το βαθμό συμμετοχής (**membership**) που έχει αυτή σε κάθε ομάδα και σε ομαδοποίηση, όπου κάθε εγγραφή κατηγοριοποιείται σε μία μόνο ομάδα (**hard clustering**).

5.4.2 ΑΛΓΟΡΙΘΜΟΙ ΟΜΑΔΟΠΟΙΗΣΗΣ

5.4.2.1 ΔΙΑΧΩΡΙΣΤΙΚΗ ΟΜΑΔΟΠΟΙΗΣΗ

Στη διαχωριστική ομαδοποίηση επιχειρείται και αποκτάται ένας μονό διαχωρισμός των δεδομένων σε ομογενείς ομάδες. Στην περίπτωση αυτή, επιχειρείται ο καθορισμένος ακέραιου αριθμού ομάδων- κατηγοριών, ο οποίος ελαχιστοποιεί μια συγκεκριμένη συνάρτηση κόστους. Η ελαχιστοποίηση αυτή μπορεί να είναι είτε τοπική (δηλαδή να αφορά ένα μέρος μόνο των εγγραφών), είτε ολική (να αφορά όλες τις εγγραφές). Ο αλγόριθμος σε αυτή την περίπτωση επαναλαμβάνεται πολλές φορές με διαφορετικές αρχικές συνθήκες, έτσι ώστε τα καλύτερα αποτελέσματα από όλες τις επαναλήψεις είναι τελικά το τελικό προϊόν της ομαδοποίησης (**Pompe, P., Feelders, A., 1997**)

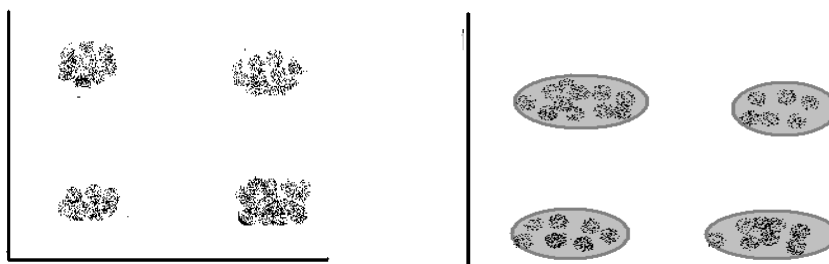
Το πιο σύνηθες μέτρο/ κριτήριο ομοιότητας των εγγραφών μιας ομάδας που χρησιμοποιείται στη διαχωριστική ομαδοποίηση είναι το κριτήριο του τετραγωνικού σφάλματος (**square error criterion**).

Οι αλγόριθμοι διαχωριστικής ομαδοποίησης εφαρμόζονται κυρίως σε αριθμητικά δεδομένα. Παρ' όλα αυτά υπάρχουν και αλγόριθμοι διαχωριστικής ομαδοποίησης που εφαρμόζονται σε κατηγορικά δεδομένα. Ένας από τους πιο γνωστούς αλγόριθμους διαχωριστικής ομαδοποίησης που εφαρμόζεται σε

κατηγορικά δεδομένα είναι ο **K- modes**, ο οποίος αποτελεί μία παραλλαγή του **K- means**. Ο αλγόριθμος **K- modes** χρησιμοποιεί την ίδια τεχνική με τον **K- means** για να ανακαλύψει τις ομάδες μέσα σε μία βάση δεδομένων, αλλά υιοθετεί νέες έννοιες για να χειρισθεί τα κατηγορικά δεδομένα. Σε αυτή την περίπτωση, τα κέντρα των ομάδων αντικαθίστανται με μέσους (**modes**), οι οποίοι αντιπροσωπεύουν ένα διαφορετικό μέτρο ανομοιότητας από αυτό περιγραφικό ήδη. Το μέτρο ανομοιότητας ορίζεται ως οι συνολικές διαφορές των αντίστοιχων κατηγορικών χαρακτηριστικών των εγγραφών μεταξύ τους και δίνει μεγαλύτερη έμφαση σε όχι συχνά εμφανιζόμενες κατηγορίες. Όσο μικρότερο είναι το πλήθος των διαφορών, τόσο περισσότερο όμοιες είναι οι έγγραφες μεταξύ τους.

Οι αλγόριθμοι διαχωριστικής ομαδοποίησης έχει αποδειχθεί ότι δεν μπορούν να χειρισθούν το θόρυβο στα δεδομένα, επηρεάζονται από απομονωμένες τιμές (**outliers**), όπως επίσης δεν μπορούν να ανακαλύψουν ομάδες με μη κυρτό σχήμα.

Στο διάγραμμα 2 φαίνεται μία διαχωριστική ομαδοποίηση δεδομένων σε έναν διδιάστατο χώρο.



5.4.2.2 ΙΕΡΑΡΧΙΚΗ ΟΜΑΔΟΠΟΙΗΣΗ

Η ιεραρχική ομαδοποίηση είτε ενώνει μικρές ομάδες σε μεγαλύτερες, είτε χωρίζει μεγάλες ομάδες σε μικρότερες. Η λειτουργεί ενός αλγόριθμου ιεραρχικής ομαδοποίησης απεικονίζεται δισδιάστατα και παράγει ένα δενδρόγραμμα, το οποίο αντιπροσωπεύει την ομαδοποίηση των προτύπων και τα επίπεδα ομοιότητας και ανομοιότητας, στα οποία η ομαδοποίηση αλλάζει. Το δενδρόγραμμα λοιπόν, χωρίζεται σε διάφορα επίπεδα, έτσι ώστε να παράγει διαφορετική ομαδοποίηση των δεδομένων.

Οι ιεραρχικοί κανόνες ομαδοποίησης, συμφωνά με την μέθοδο που χρησιμοποιούν για να σχηματίσουν ομάδες μπορούν να χωριστούν ως ακολούθως :

- I. Συγκεντρωτικοί αλγόριθμοι (**agglomerative algorithms**) : Οι αλγόριθμοι αυτοί παράγουν μια σειρά από ομαδοποιήσεις, οι οποίες μειώνουν τον αριθμό των ομάδων σε κάθε βήμα του αλγόριθμου.
- II. Διαιρετικοί αλγόριθμοι (**divisive algorithms**) : Οι αλγόριθμοι αυτοί παράγουν μια σειρά από ομαδοποιήσεις, οι οποίες αυξάνουν τον αριθμό των ομάδων σε κάθε βήμα του αλγόριθμου.

Η βασική διαδικασία της ιεραρχικής ομαδοποίησης έχει ως ακολούθως :

1. Κάθε εγγραφή κατατάσσεται σε μία ομάδα. Έτσι, εάν υπάρχουν N εγγραφές υπάρχουν και N ομάδες. Οι αποστάσεις (ομοιότητες) κάθε ομάδας είναι οι ίδιες οι αποστάσεις (ομοιότητες) των στοιχείων που περιέχουν.
2. Οι περισσότερες ίδιες ομάδες (ζευγάρι) ομαδοποιούνται σε μία ομάδα, έτσι ώστε υπάρχει πλέον μία ομάδα λιγότερη.
3. Υπολογίζονται οι αποστάσεις (ομοιότητες) μεταξύ της νέας ομάδας και κάθε μιας από τις παλαιότερες ομάδες.

4. Γίνεται επανάληψη των βημάτων 2 και 3, μέχρι ότου όλες οι εγγραφές ανήκουν πλέον σε μία ομάδα μεγέθους N .

Εφόσον η διαδικασία τερματιστεί, ο χρήστης έχει στη διάθεση του ένα πλήρες δενδρόγραμμα, το οποίο του δίνει την δυνατότητα να επιλέξει το επίπεδο, στο οποίο θέλει να σταματήσει την ομαδοποίηση.

Το τρίτο στάδιο μπορεί να πραγματοποιηθεί με διαφόρους τρόπους. Έτσι, οι αλγόριθμοι ιεραρχικής ομαδοποίησης μπορούν να χωριστούν σε ομαδοποίηση μονή σύνδεσης (**single-link**), ολικής σύνδεσης (**complete link**) και μέσης σύνδεσης (**average-link**). Στην πρώτη κατηγορία η απόσταση μεταξύ δύο ομάδων είναι ίση με τη μικρότερη απόσταση από οποιαδήποτε εγγραφή μιας ομάδας με οποιαδήποτε εγγραφή άλλης ομάδας. Στη δεύτερη κατηγορία η απόσταση μεταξύ δύο ομάδων είναι ίση με τη μεγαλύτερη απόσταση από οποιαδήποτε εγγραφή μιας ομάδας με οποιαδήποτε εγγραφή άλλης ομάδας. Τέλος, στην τρίτη κατηγορία η απόσταση μεταξύ δύο ομάδων είναι ίση με τη μέση απόσταση από οποιαδήποτε εγγραφή μιας ομάδας με οποιαδήποτε εγγραφή άλλης ομάδας.

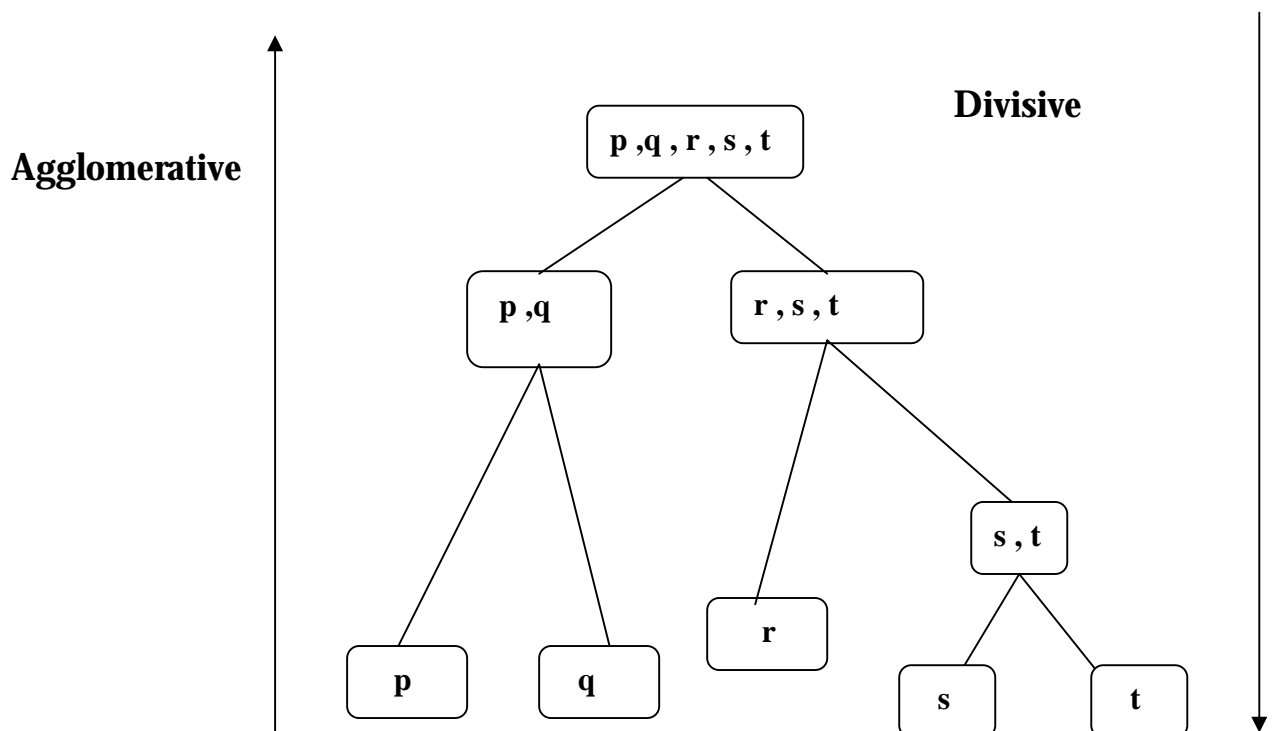
Οι αλγόριθμοι που ανήκουν στην παραπάνω κατηγορία ονομάζονται συγκεντρωτικοί αλγόριθμοι, διότι ενώνουν ομάδες σε κάθε μία επανάληψη. Όπως ήδη αναφέρθηκε, υπάρχουν και οι διαιρετικοί αλγόριθμοι ιεραρχικής ομαδοποίησης, οι οποίοι όμως δεν είναι ιδιαίτερα διαθέσιμοι και δεν έχουν χρησιμοποιηθεί ευρέως.

Άλλοι αλγόριθμοι ιεραρχικής ομαδοποίησης περιλαμβάνουν τον **CURE**, ο οποίος χρησιμοποιεί τεχνικές τυχαίας δειγματοληψίας και διαχωριστική ομαδοποίηση, προκειμένου να κατατάξει τα δεδομένα σε ομάδες που περιλαμβάνουν ένα συγκεκριμένο αριθμό εγγραφών, οι οποίες μετακινούνται προς το κέντρο των ομάδων, βάσει ενός κλάσματος, καθώς και τον **ROCK**, ο οποίος εφαρμόζεται σε κατηγορικά δεδομένα και εισάγει δύο νέες έννοιες, τις γειτονικές εγγραφές και τις συνδέσεις μεταξύ δύο εγγραφών.

Τέλος, ο αλγόριθμος **OPTICS** σε αντίθεση με τις υπόλοιπες μεθόδους, δεν παράγει απλώς μία ομαδοποίηση των δεδομένων, αλλά παράγει μια αυξητική κατάταξη των δεδομένων, η οποία αντιπροσωπεύει τη δομή της ομαδοποίησης βάσει της πυκνότητας της (**Shin, K., Lee, T., Kim, H., 2005**).

Οι αλγόριθμοι ιεραρχικής ομαδοποίησης γενικά είναι πιο αποτελεσματικοί στη διαχείριση του θορύβου από ό,τι οι διαχωριστικοί αλγόριθμοι.

Στο διάγραμμα **3** φαίνεται μια ιεραρχική ομαδοποίηση δεδομένων, τόσο σε συγκεντρωτικό, όσο και σε διαιρετικό επίπεδο.



5.4.2.3 ΟΜΑΔΟΠΟΙΗΣΗ ΒΑΣΗ ΣΥΝΘΗΚΩΝ ΠΥΚΝΟΤΗΤΑΣ

Οι αλγόριθμοι που ανήκουν σε αυτή την κατηγορία τυπικά θεωρούν τις ομάδες που ομαδοποιούν ως περιοχές μεγάλης πυκνότητας στο χώρο των δεδομένων (**data space**), οι οποίες διαχωρίζονται από περιοχές μικρής πυκνότητας.

Δύο γνωστοί αλγόριθμοι που ανήκουν σε αυτή την κατηγορία είναι ο **DBSCAN** και ο **DENCLUE**. Στον πρώτο, για κάθε εγγραφή σε μία ομάδα, μία περιοχή με δεδομένη ακτίνα, πρέπει να περιλαμβάνει έναν ελάχιστο αριθμό εγγραφών, δηλαδή η πυκνότητα σε μία ομάδα πρέπει να υπερβαίνει κάποιο κατώφλι. Στο δεύτερο, το σύνολο των εγγραφών μοντελοποιείται αναλυτικά ως ένα άθροισμα συναρτήσεων επιρροής των εγγραφών. Η συνάρτηση αυτή μπορεί να θεωρηθεί ότι περιγράφει την επίδραση που έχει μία εγγραφή στην περιοχή, στην οποία ανήκει. Στη συνέχεια, οι εγγραφές κατατάσσονται σε ομάδες βάσει αυτών των συναρτήσεων πυκνότητας, οι οποίες είναι τοπικά μέγιστα της συνολικής συνάρτησης πυκνότητας. Τέλος, στους αλγόριθμους **x- windows**, **k- windows** και **z- windows** χρησιμοποιείται μία παραθυρική τεχνική, προκειμένου να μειωθεί ο αριθμός των υπό εξέταση προτύπων, ενώ στον αλγόριθμο **z- windows** ο αριθμός των αρχικών ομάδων καθορίζεται ενδογενώς από το σύστημα.

Οι αλγόριθμοι που ανήκουν σε αυτή την κατηγορία μπορούν να χειρισθούν δεδομένα με τυχαίο σχήμα στο χώρο (π.χ. ελλειψοειδή, κυλινδρικά κλπ.), καθώς και ομάδες με διαφορετικά μεγέθη, όπως επίσης και το θόρυβο στα δεδομένα.

5.4.2.4 ΟΜΑΔΟΠΟΙΗΣΗ ΒΑΣΗ ΠΛΕΓΜΑΤΩΝ

Οι αλγόριθμοι που ανήκουν σε αυτή την κατηγορία διαιρούν το χώρο σε ένα πεπερασμένο αριθμό κελιών, για να ομαδοποιήσουν τα δεδομένα, εφαρμόζοντας τις μεθόδους ξεχωριστά σε αυτούς τους χώρους.

Τρεις ιδιαίτερα γνωστοί αλγόριθμοι που ανήκουν σε αυτή την κατηγορία είναι ο **STING**, ο **WaveCuster** και ο **CLIQUE**. Ο **STING** χωρίζει το χώρο σε ορθογώνια κελιά, χρησιμοποιώντας μια ιεραρχική δομή. Στη συνέχεια υπολογίζει στατιστικές παραμέτρους (όπως μέσο, διάμεσο, τύπο κατανομής κλπ.) όλων των χαρακτηριστικών όλων των εγγραφών σε αυτούς τους χώρους. Ο αλγόριθμος **WaveCuster** κατ' αρχήν ομαδοποιεί τα δεδομένα βάσει μίας πολυδιάστατης δομής, έτσι ώστε κάθε κελί βάσει των εγγραφών που εμπεριέχονται σε αυτό. Τέλος, ο αλγόριθμος **CLIQUE** ανακαλύπτει πυκνές ομάδες σε υποχώρους μέγιστων διαστάσεων και δημιουργεί περιγραφές ομάδων βάσει **DNF** εκφράσεων. Γενικά οι αλγόριθμοι αυτοί εφαρμόζονται αποτελεσματικά σε μεγάλες βάσεις δεδομένων.

5.4.2.5 ΟΜΑΔΟΠΟΙΗΣΗ FUZZY

Οι μεθοδολογίες που αναπτύχθηκαν ως τώρα αντιστοιχούν σε ομαδοποίηση που παράγει ομάδες, όπου κάθε στοιχείο συμπεριλαμβάνεται σε μια μόνο ομάδα. Παρ' όλα αυτά η αβεβαιότητα που προέρχεται από την ομαδοποίηση και το γεγονός ότι τα όρια μεταξύ των ομάδων δεν είναι πάντα απόλυτα ορισμένα, οδηγεί στην ομαδοποίηση **fuzzy**, όπου στοιχεία (εγγραφές) μπορεί να ανήκουν σε περισσότερες από μία ομάδες, βάσει μίας συνάρτησης βαθμού συμμετοχής. Η ομαδοποίηση **fuzzy** έχει επεκταθεί, ώστε να μπορεί να εφαρμοσθεί και σε κατηγορικά δεδομένα (**Sikonja M. And Kononenko I., 1997**).

5.4.2.6 ΟΜΑΔΟΠΟΙΗΣΗ ΒΑΣΗ ΚΑΤΑΝΟΜΩΝ GAUSSIAN

Μια άλλη μεθοδολογία που χρησιμοποιείται για την ομαδοποίηση δεδομένων βασίζεται στην μοντελοποίηση, η οποία χρησιμοποιεί συγκεκριμένα μοντέλα ως ομάδες και επιχειρεί να μεγιστοποιήσει τη συνάφεια μεταξύ των δεδομένων και του μοντέλου. Αυτό σημαίνει ότι κάθε ομάδα μπορεί μαθηματικά να παρουσιασθεί σαν μία παραμετρική κατανομή, η οποία είναι είτε συνεχής (π.χ. **Gaussian**) , είτε διακριτή (π.χ. **Poisson**). Έτσι, το σύνολο των εγγραφών μοντελοποιείται σαν μία μίξη (**mixture**) αυτών των κατανομών. Η πιο ευρέως χρησιμοποιούμενη μεθοδολογία βασίζεται σε μία μίξη κατανομών **Gaussian**, όπου κάθε ομάδα αντιπροσωπεύεται από μία κατανομή **Gaussian**, με πρότυπο το βαρύκεντρό της.

5.4.3 ΔΙΑΦΟΡΑ ΘΕΜΑΤΑ

Οι τεχνικές ομαδοποίησης παρουσιάζουν διάφορα προβλήματα, ανάμεσα στα οποία και τα εξής :

- Οι υπάρχουσες τεχνικές ομαδοποίησης δεν είναι απόλυτα ικανές να διαχειρισθούν ένα μεγάλο αριθμό χαρακτηριστικών, λόγω της πολυπλοκότητας του χρόνου που εμπλέκεται κατά τη διαδικασία ομαδοποίησης.
- Η αποτελεσματικότητα της μεθόδου εξαρτάται από τον ορισμό της «απόστασης» μεταξύ δύο σημείων (σε περίπτωση ομαδοποίησης βάσει απόστασης δύο σημείων).
- Εάν δεν υπάρχει ένα προφανές μέτρο γειτνίασης, τότε είναι ο ίδιος ο χρήστης που πρέπει να το ορίσει, το οποίο δεν είναι πάντα εύκολο, λόγω των πολυδιάστατων χώρων των βάσεων δεδομένων.

- Το αποτέλεσμα της ομαδοποίησης μπορεί να ερμηνευτεί με διάφορους τρόπους, και όχι πάντα με έναν συγκεκριμένο.

ΣΥΜΠΕΡΑΣΜΑΤΑ

Από την προηγούμενη ανάλυση καθίσταται σαφές ότι τα πληροφοριακά συστήματα αποτελούν αναπόσπαστο κομμάτι μιας σύγχρονης επιχείρησης. Σημαντική ώθηση στην ανάπτυξη των επιχειρήσεων στο σύγχρονο ανταγωνιστικό περιβάλλον, με τον τεράστιο όγκο διαθέσιμων πληροφοριών, έδωσε ο τομέας της επιχειρηματικής νοημοσύνης που παρακάμπτοντας την υποκειμενική ανθρώπινη χρήση, παρέχει την δυνατότητα λήψης γρήγορων και ασφαλών αποφάσεων ή την υποβοήθηση της απόφασης. Στην χώρα μας συστήματα επιχειρηματικής νοημοσύνης για την λήψη ή την υποβοήθηση αποφάσεων, συναντώνται μόνο σε μεγάλα χρηματοπιστωτικά ιδρύματα, κυρίως λόγω του μεγάλου τους κόστους. Παρόλα αυτά, στα επόμενα χρόνια, η αναμενόμενη αύξηση του όγκου το πληροφοριών που θα εισρέουν σε κάθε επιχείρηση, σε συνδιασμό με την αναμενόμενη μείωση του κόστους των συστημάτων λήψης αποφάσεων και την αύξηση του ανταγωνιστικού πλαισίου μέσα στο οποίο κινούνται οι επιχειρήσεις, αναμένεται να καταστήσουν τα συστήματα επιχειρηματικής νοημοσύνης απαραίτητα εργαλεία και για τις επιχειρήσεις μεσαίου βελινεκούς.

BIBΛΙΟΓΡΑΦΙΑ

- Aha, D. (1997), *Lazy Learning*, Dordrecht: Kluwer Academic Publishers.
- Altman, E.L. (1993). *Corporate Financial Distress and Bankruptcy*. John Wiley and Sons.
- Cohen, W. (1995), "Fast Effective Rule Induction", *Proceeding of international Conference on Machine Learning 1995*, pp. 115-123.
- Dimitras, A. I. Slowinski, R. Susmaga, R. and Zopounidis, C. (1999). Business failure prediction using rough sets. *European Journal of Operational Research* 114: 263-280
- Domingos, P. and Pazzani, M. (1997), "On the optimality of the simple Bayesian classifier under zero-one loss. *Machine Learning*", Vol.29, pp.103-130.
- Doumpos, M., Zopounidis, C. (1999), A Multicriteria Discrimination Method for the Prediction of Financial Distress: the case of Greece, *Multinational Finance Journal*, 1999, vol. 3, no. 2, pp. 71-101.
- Freund, Y. and Schapire, R.E. (1997). A decision- theoretic generalization of online learning and an application to boosting. *J. Comput. System Sciences* 55.
- Friedman, J., Hastie, T. and Tibshirani, R. (2000), Additive logistic regression: a statistical view of boosting, *Ann. Statist.* 28, no. 2 (2000), 337-407.
- Japkowicz N. and Stephen, S. (2002), The Class Imbalance Problem: A Systematic Study *Intelligent Data Analysis*, Volume 6, Number 5.
- Kotsiantis, S., Zaharakis, I., Tampakas, V., Pintelas, P. (2004), On Constructing a Financial Decision Support System, *Proceedings of International Conference on Enterprise Systems and Accounting 2004*, September 3- 4, Thessaloniki, Greece, pp 319-331.

Mckee, T., Greenstein, M. (200), Predicting Bankruptcy Using Recursive Partitioning and a Realistically Proportioned Data Set, *Journal of Forecasting*, 19, 219- 230.

Mitchell, T. (1997), *Machine Learning*, McGraw Hill.

Morris, R., 1998. *Early Warning Indicators of Corporate Failure: A Critical Review of Previous Research and Further Empirical Evidence*. Ashgate Publishing Company.

Negakis, C. (1995), "Robustness of Greek Business failure prediction models", *International Review of Economics and Business*, Vol. 42 no. 3, pp. 203- 215.

Papoulias, C., and Theodossiou, R. (1992), Analysis and modeling of recent business failures in Greece. *Managerial and Decision Economics* 13(2): 163-169.

Park, C., Han, I. (2002), A Case- Based Reasoning with the Feature Weights Derived by Analytic Hierarchy Process for Bankruptcy Prediction. *Expert systems with Applications* 23(3), 225-264.

Platt, J. (1999), Using sparseness and analytic QP to speed training of support vector machines. In M. S. Kearns, S.A. Solla, & D.A Cohn (Eds.), *Advances in neural information processing systems* 11. MA: MIT Press.

Pompe, P., Feelderrs, A. (1997), Using Machine Learning, Neural Networks, and Statistics to Predict Corporate Bankruptcy. *Microcomputers in Civil Engineering* 12, 267-276.

Quilan, J. R. (1993), *C4.5: Programs for machine learning*, Morgan Kaufmann, San Francisco.

Sarkar, S., and R. S. Sriram, (2001), Bayesian Models for Early Warning of Bank Failures. *Management Science*, 47(11): 1457-1475.

Shin, K., Lee, T., Kim, H., (2005), An application of support vector machines in bankruptcy prediction model, *Expert systems with Applications*, volume 28, pp. 127-135.

Sikonja M. and Kononenko I. (1997), An adaptation of Relief for attribute estimation in regression, Proceedings of the Fourteenth International Conference (ICML'97), ed., Dough Fisher, pp. 296-304. Morgan Kaufmann Publishers.

Thomaidis, N., Gounias, G., Zopounidis, C., (1999). A fuzzy rule based learning method for corporate bankruptcy prediction, ACAI 99, Chania, Greece.

Vranas, A. S. (1992), The significance of financial characteristics in predicting business failure: an analysis in the Greek context. Foundation of Computing and Decision Sciences 17 (4): 257-275.

Witten, I. and Frank E. (2005), Data mining: Practical Machine Learning Tools and Techniques with java Implementations, 2nd Edition, Morgan Kaufmann, San Francisco, 2005.

Zhang, G., Hu, M. Y., Patuwo, B.E., Indrro, D. C. (1999), Artificial Neural Networks in Bankruptcy Prediction: General Framework and Cross-Validation analysis. European Journal of Operationnal Research 116, 16-32.

Νόμος 3588/2007 «Πτωχευτικός Κώδικας» ΦΕΚΑ' 153/10.7.2007