

Η διαδικασία απόκτησης της γνώσης απλουστεύεται γιατί τις περισσότερες φορές η γνώση των παλιών περιπτώσεων υπάρχει ήδη σε παλαιότερα έγγραφα μιας εταιρίας ή ακόμα καλύτερα σε βάσεις δεδομένων. Εάν δεν υπάρχει βιβλιοθήκη παλιών περιπτώσεων, τότε κάποιος ειδικός καλείται να τη δημιουργήσει από τις εμπειρίες του. Η περίπτωση αυτή βέβαια απαιτεί πολύ μεγαλύτερο κόπο εκ μέρους του ειδικού από μια απλή συμμετοχή του στη διαδικασία εκμείωσης γνώσης όπως γίνεται στα έμπειρα συστήματα.

Η βιβλιοθήκη των περιπτώσεων μπορεί να αποτελείται από τη συλλογική εμπειρία ενός οργανισμού ή μιας εταιρίας και όχι ενός μόνο ειδικού. Στο πλαίσιο αυτό κινείται η έρευνα στον τομέα της *διαχείρισης της γνώσης (knowledge management)* που είναι μια διαδικασία μετατροπής της γνώσης από τις πηγές στις οποίες είναι διαθέσιμη μέσα σε κάποιον οργανισμό σε κάποια ηλεκτρονική μορφή αναπαράστασης γνώσης, όπως ηλεκτρονικά έγγραφα ή ακόμα και κανόνες, και στη συνέχεια σύνδεσης αυτής της γνώσης με τα στελέχη του οργανισμού. Η διαχείριση της γνώσης περιλαμβάνει εργαλεία και μεθόδους για τη δημιουργία, ανάκληση και επαναχρησιμοποίηση της γνώσης.

Το κυριότερο μειονέκτημα της συλλογιστικής των περιπτώσεων είναι το υπολογιστικό κόστος της αναζήτησης στη βιβλιοθήκη των περιπτώσεων. Για το λόγο αυτό είναι σημαντικό το ζήτημα της "έξυπνης" δεικτοδότησης, ώστε να επιταχύνεται η αναζήτηση. Άλλα σημαντικά ζητήματα που επηρεάζουν την απόδοση και την ποιότητα των λύσεων που παρέχει ένα τέτοιο σύστημα είναι η δομή της βιβλιοθήκης καθώς και η ποιότητα και ποσότητα των περιπτώσεων που βρίσκονται στη βιβλιοθήκη. Τα παραπάνω ζητήματα δεν αποτελούν πρόβλημα μόνο για τα συστήματα που στηρίζονται στη συλλογιστική των περιπτώσεων αλλά για κάθε πληροφοριακό σύστημα, οπότε μπορεί να ξεπεραστούν αν αντιμετωπιστούν με ιδιαίτερη προσοχή. Τέλος, πολλές φορές εμφανίζονται δυσκολίες στην προσαρμογή της λύσης στην τρέχουσα κατάσταση, ιδιαίτερα αν δεν υπάρχουν αριθμητικές μέθοδοι αλλά απαιτούνται εμπειρικές σχέσεις.

23.3.2 Μελέτη Περίπτωσης

Στην ενότητα αυτή παρουσιάζεται η μελέτη περίπτωσης του συστήματος PAS (*Property Appraisal System*) το οποίο *βασίζόταν στη συλλογιστική των περιπτώσεων για τον αυτόματο προσδιορισμό της αξίας μιας ακίνητης ιδιοκτησίας*. Το σύστημα PAS χρησιμοποιούσε κάποια συγκεκριμένα χαρακτηριστικά στοιχεία της ιδιοκτησίας (Πίνακας 23.3) για σύγκριση με τη βιβλιοθήκη περιπτώσεων. Το πρώτο βήμα στη διαδικασία ήταν η ανάκληση της κατάλληλης περίπτωσης. Το σύστημα επικεντρωνόταν στις πιο πρόσφατες αγοραπωλησίες της βιβλιοθήκης, ανακαλούσε τις 10 πιο συναφείς περιπτώσεις, τις οποίες βαθμολογούσε και ταξινομούσε κατά φθίνουσα σειρά ομοιότητας.

Για τη βαθμολόγηση των περιπτώσεων έπρεπε να καθοριστούν τα ακόλουθα:

- Τα βάρη ή η σπουδαιότητα κάθε χαρακτηριστικού βάσει του οποίου γίνεται η σύγκριση. Για παράδειγμα, το εμβαδόν παίζει πιο σπουδαίο ρόλο από τον όροφο.
- Ο τρόπος με τον οποίο θα βαθμολογούνται οι διαφορές στη σύγκριση μεταξύ των χαρακτηριστικών. Για παράδειγμα, αν η ηλικία του σπιτιού της τρέχουσας περίπτωσης είναι 22 χρόνια σε σχέση με τα 20 χρόνια που είναι η ηλικία ενός σπιτιού που βρίσκεται στη βιβλιοθήκη περιπτώσεων, τότε η μικρή αυτή διαφορά δεν παίζει μεγάλο ρόλο.

Πίνακας 23.3: Παράδειγμα καθορισμού αξίας ακίνητης περιουσίας.

Χαρακτηριστικό	Βάρος	Τρόπος βαθμολόγησης διαφορών	Τρόπος προσαρμογής περίπτωσης
Καθαρό εμβαδόν (<i>min=35, max=210</i>)	0.9	Απόλυτη Ποσοστιαία Διαφορά (ΑΠΔ)	Διαφορά · Κατασκευαστική τιμή m ²
Αριθμός δωματίων (<i>min=1, max=5</i>)	0.8	ΑΠΔ	Διαφορά · 6,000 €
Αριθμός τουαλετών (<i>min=1, max=3</i>)	0.5	ΑΠΔ	Διαφορά · 3,000 €
Αρχιτεκτονικός ρυθμός (<i>πολυκατοικία=0, μεζονέτα=0.8, μονοκατοικία=1</i>)	1.0	Απόλυτη Διαφορά (ΑΔ)	Διαφορά · 30% (της τιμής του σπιτιού της βάσης δεδομένων)
Ηλικία οικήματος (σε χρόνια) (<i>min=0, max=10</i>)	0.7	ΑΠΔ	Διαφορά · 2% (της τιμής του σπιτιού της βάσης)
Θέση (περιοχή - γειτονιά) (<i>min=0, max=25</i>)	0.8	ΑΠΔ - Απόσταση από το κέντρο της πόλης	Διαφορά αντικειμενικής (ανά περιοχή) αξίας m ² · Εμβαδόν
Ημερομηνία αγοραπωλησίας (<i>min=0, max=6</i>)	0.6	ΑΠΔ - Χρονική διαφορά/τριετία	Διαφορά (σε χρόνια) · 3% (της τιμής του σπιτιού της βάσης δεδομένων)
Τύπος ψύξης (<i>aircondition=1, ανεμιστήρας=0.5, καθόλου=0</i>)	0.2	ΑΔ	Διαφορά · 1% (της τιμής του σπιτιού της βάσης δεδομένων)
Τύπος θέρμανσης (<i>ατομική αερίου=1, κεντρική πετρελαίου=0.5, θερμοσυσσωρευση=0.25, καθόλου=0</i>)	0.7	ΑΔ	Διαφορά · 4% (της τιμής του σπιτιού της βάσης δεδομένων)
Τύπος parking (<i>υπόγειο=1, πυλωτή=0.75, ανοιχτό-ιδιωτικό=0.5, καθόλου=0</i>)	0.3	ΑΔ	Διαφορά · 10% (της τιμής του σπιτιού της βάσης δεδομένων)
Μέγεθος οικοπέδου (<i>min=500, max=4000</i>)	0.2	ΑΠΔ	Διαφορά · 300 €
Ύπαρξη πισίνας (<i>ναι=1, όχι=0</i>)	0.1	ΑΔ	Διαφορά · 25% (της τιμής του σπιτιού της βάσης δεδομένων)

Στο συγκεκριμένο παράδειγμα υπάρχουν δύο τρόποι υπολογισμού των διαφορών, η απόλυτη ποσοστιαία διαφορά (ΑΠΔ) για τα χαρακτηριστικά των οποίων η τιμή δεν βρίσκεται στο διάστημα [0,1] και η απόλυτη διαφορά (ΑΔ) για όσα χαρακτηριστικά έχουν τιμή στο παραπάνω διάστημα. Για παράδειγμα, το καθαρό εμβαδόν μετριέται σε τετραγωνικά μέτρα (m²) και οι τιμές του είναι εκτός του διαστήματος [0,1]. Ο υπολογισμός της διαφοράς $ADiff_p$ για ένα τέτοιου τύπου χαρακτηριστικό p μεταξύ δύο σπιτιών new (για το νέο σπίτι) και j (για το σπίτι της βιβλιοθήκης) δίνεται από τη σχέση:

$$ADiff_p^{new,j} = \frac{|val_p^{new} - val_p^j|}{\max_i val_p^i - \min_i val_p^i}$$

όπου, val_p^x είναι η τιμή του χαρακτηριστικού p για το σπίτι x . Στον παρονομαστή της παραπάνω εξίσωσης υπάρχει η μέγιστη διαφορά που μπορεί να υπάρξει στην βιβλιοθήκη για το συγκεκριμένο χαρακτηριστικό, μεταξύ της μέγιστης και της ελάχιστης τιμής, με σκοπό την κανονικοποίηση της διαφοράς $ADiff$ στο διάστημα $[0,1]$. Για παράδειγμα, αν ένα σπίτι στην βιβλιοθήκη έχει εμβαδόν 120 m^2 , το υπό-εκτίμηση σπίτι έχει εμβαδόν 100 m^2 , το σπίτι με το μικρότερο εμβαδόν στην βάση είναι 35 m^2 , ενώ αυτό με το μεγαλύτερο εμβαδόν 210 m^2 , τότε η διαφορά του υπό εκτίμηση σπιτιού από το σπίτι στην βιβλιοθήκη υπολογίζεται σε $|100-120|/(210-35)=0.114$.

Όσον αφορά στα χαρακτηριστικά τα οποία ήδη βρίσκονται στο διάστημα $[0,1]$, η διαφορά υπολογίζεται από τον τύπο:

$$ADiff_p^{new,j} = |val_p^{new} - val_p^j|$$

Στο συγκεκριμένο παράδειγμα τέτοιου είδους χαρακτηριστικά είναι αυτά των οποίων οι τιμές δεν είναι καταρχήν αριθμητικές αλλά συμβολικές, όπως είναι ο τύπος θέρμανσης και ψύξης, οι οποίες αντιστοιχούνται σε αριθμούς στο διάστημα $[0,1]$, με τέτοιον τρόπο ώστε να υπάρχει πλήρης διάταξη των τιμών από την "χειρότερη" (0) στην "καλύτερη" (1). Για παράδειγμα, εάν κάποιο σπίτι δεν έχει καθόλου θέρμανση τότε η τιμή στο συγκεκριμένο χαρακτηριστικό αντιστοιχεί στο 0, εάν έχει κεντρική θέρμανση με πετρέλαιο αντιστοιχεί στο 0.5, ενώ αν έχει ατομική θέρμανση αερίου αντιστοιχεί στο 1. Αν, λοιπόν, ένα σπίτι στην βιβλιοθήκη έχει κεντρική θέρμανση με πετρέλαιο και το υπό-εκτίμηση σπίτι έχει ατομική θέρμανση αερίου, τότε η διαφορά του υπό εκτίμηση σπιτιού από το σπίτι στην βιβλιοθήκη υπολογίζεται σε $|1-0.5|=0.5$.

Η ομοιότητα των δύο σπιτιών ως προς οποιοδήποτε χαρακτηριστικό είναι πάντα το συμπλήρωμα της διαφοράς ως προς 1:

$$Sim_p^{new,j} = 1 - ADiff_p^{new,j}$$

Συνεπώς στα προηγούμενα παραδείγματα, η ομοιότητα του υπό εκτίμηση σπιτιού με το σπίτι στην βιβλιοθήκη ως προς το εμβαδόν υπολογίζεται σε 0.886, ενώ ως προς την θέρμανση είναι 0.5.

Η ομοιότητα σε κάθε χαρακτηριστικό πολλαπλασιάζεται με το βάρος του χαρακτηριστικού, το οποίο κυμαίνεται στο διάστημα $[0,1]$ προκειμένου να εκτιμηθεί η επίπτωσή του στη συνολική ομοιότητα μεταξύ των δύο σπιτιών:

$$WSim_p^{new,j} = weight_p \cdot Sim_p^{new,j}$$

Στα προηγούμενα παραδείγματα, η ομοιότητα με βάρος του υπό εκτίμηση σπιτιού με το σπίτι στην βιβλιοθήκη ως προς το χαρακτηριστικό του εμβαδού υπολογίζεται σε $0.886 \cdot 0.9 = 0.7974$, ενώ ως προς την θέρμανση είναι $0.5 \cdot 0.7 = 0.35$.

Τέλος, η συνολική ομοιότητα δύο σπιτιών υπολογίζεται ως το άθροισμα των ομοιοτήτων με βάρη όλων των χαρακτηριστικών p και για να είναι κανονικοποιημένη στο διάστημα $[0,1]$ το άθροισμα διαιρείται με τον συνολικό αριθμό n των χαρακτηριστικών:

$$Sim^{new,j} = \frac{\sum_{p=1}^n WSim_p^{new,j}}{n} = \frac{1}{n} \cdot \sum_{p=1}^n weight_p \cdot Sim_p^{new,j}$$