

# Καινοτόμος έρευνα στη βάση των αναγκών

---

Παραδοτέο 3.4 (έκδοση 1.1)

9/5/2019



## Πίνακας Περιεχομένων

1	Εισαγωγή .....	3
2	Βασικές Λειτουργίες.....	4
3	Περιγραφή Λειτουργιών ΙΤΕ.....	<b>Error! Bookmark not defined.</b>
3.1	Σημειογραφία.....	5
3.2	Συμπιεστική δειγματοληψία .....	6
3.3	Παρακολούθηση συσχετίσεων .....	8
3.4	Ανακατασκευή χαμένων μετρήσεων .....	9
3.5	Αύξηση της χρονικής ανάλυσης.....	10
3.6	Αναγνώριση μη φυσιολογικής συμπεριφοράς.....	11
4	Περιγραφή Λειτουργιών ΚΟΙΟΣ.....	<b>Error! Bookmark not defined.</b>
4.1	Εκτίμηση υδραυλικής κατάστασης δικτύου .....	12
4.2	Έλεγχος πίεσης δικτύου με χρονοδρομολόγηση βαλβίδων μείωσης πίεσης.....	12
4.3	Αξιολόγηση ρίσκου σπάσιμου σωλήνων.....	13
4.4	Ανίχνευση διαρροών.....	14
4.5	Εκτίμηση ποιοτικής κατάστασης του νερού στο δίκτυο .....	16
4.6	Ανίχνευση ανωμαλιών στην συγκέντρωση χλωρίου .....	18
5	Σύνοψη.....	19

# 1 Εισαγωγή

Το παραδοτέο 3.4 έχει ως στόχο την παρουσίαση και περιγραφή όλων των αποτελεσμάτων της καινοτόμου έρευνας που διεξήχθη από το ΙΤΕ και το ΚΟΙΟΣ για τις ανάγκες της πράξης, τα οποία θα ενσωματωθούν στην πλατφόρμα. Τα αποτελέσματα των ερευνών είναι υπό δημοσίευση σε επιστημονικά περιοδικά και συνέδρια (βλ. Παραδοτέο 2.5). Οι βασικές λειτουργίες που θα αναπτυχθούν στην συνέχεια αφορούν τις ανάγκες όλων των εταιρών και γενικότερα του έργου SmartWater2020 και πιο συγκεκριμένα:

- Ανίχνευση διαρροών με την χρήση αισθητήρων πίεσης
- Ανίχνευση μόλυνσης του νερού και εκτίμηση της κατάστασής του συστήματος
- Έλεγχος της πίεσης στο δίκτυο με την χρήση βαλβίδων ρύθμισης σε πραγματικό χρόνο
- την παρακολούθηση του δικτύου νερού για αναγνώριση γεγονότων μη κανονικής λειτουργίας ως προς την ποιότητα του νερού
- Την εκτίμηση ρίσκων βλάβης στο δίκτυο υδροδότησης
- Το συνδυασμό τεχνικών κατανεμημένης διαχείρισης ασύρματων δικτύων αισθητήρων/επενεργητών που λειτουργούν εντός του δικτύου υδροδότησης με τεχνικές μηχανικής μάθησης (π.χ., ανεπιτήρητη αναγνώριση επικρατέστερων στατιστικών χαρακτηριστικών), για τη σε-πραγματικό χρόνο αναγνώριση των σημαντικότερων χαρακτηριστικών που επηρεάζουν την απόδοση του δικτύου, και τον κατάλληλο συντονισμό των δικτυακών παραμέτρων
- Τον κατανεμημένο έλεγχο της τοπολογίας του δικτύου ως προς την ισχύ μετάδοσης και τον κύκλο λειτουργίας (radio duty cycle) των ενεργειακά αυτόνομων ασύρματων αισθητήριων πλατφορμών και επενεργητών, με στόχο την επιμήκυνση του χρόνου ζωής του δικτύου
- Το συγκερασμό τεχνικών συμπιεσμένης δειγματοληψίας με δικτυακούς αλγορίθμους για τη δυναμική και αυτό-αναπροσαρμοζόμενη ρύθμιση της συχνότητας δειγματοληψίας από το δίκτυο αισθητήρων μέτρησης ποιότητας του νερού, με απώτερο στόχο την εξοικονόμηση του απαιτούμενου εύρους ζώνης/καταναλισκόμενης ενέργειας, αλλά χωρίς να υποβαθμίζεται η ποιότητα ανακατασκευής του πεδίου δειγματοληψίας στο κέντρο ελέγχου
- Την ανάπτυξη δικτυακών πρωτοκόλλων και αρχιτεκτονικών που συνδυάζουν στατικούς αισθητήρες με κινούμενους υποβρύχιους αισθητήριους κόμβους, οι οποίοι διευκολύνουν την επιτόπια ανίχνευση βλαβών / δυσλειτουργιών (π.χ. διαρροές ή/και αλλοίωση ποιότητας) και την ανάπτυξη συνδυαστικών έξυπνων εφαρμογών για τη διευκόλυνση επιτόπιων ελέγχων από τεχνικούς.

Στόχος του παραδοτέου είναι η ανάλυση όλων των επιμέρους λειτουργιών και των αντίστοιχων αλγορίθμων. Επιπρόσθετα, η διαδικτυακή πλατφόρμα που δημιουργήθηκε στο πλαίσιο του SmartWater2020 θα διασυνδεθεί με τα υφιστάμενα συστήματα των οργανισμών διαχείρισης νερού, καθώς επίσης και με τα καινούργια συστήματα που έχουν αποκτήσει οι υδατοπρομήθειες. Η πλατφόρμα θα αξιοποιεί τις μεθοδολογίες που περιγράφονται στο παραδοτέο, οι οποίες έχουν αναπτυχθεί από την ερευνητική κοινότητα για την ανάλυση του μεγάλου όγκου δεδομένων που παράγεται από τα συστήματα αυτά, για την εκτίμηση της κατάστασης του συστήματος, τη διάγνωση προβλημάτων, ως επίσης και για τον βέλτιστο έλεγχο της πίεσης στο σύστημα και βασίζονται στη μελέτη των πραγματικών αναγκών των οργανισμών ύδρευσης.

## 2 Βασικές Λειτουργίες

Οι βασικές λειτουργίες της πλατφόρμας του SmartWater2020 μπορούν να ταξινομηθούν σε δυο κατηγορίες: α) λειτουργίες που αφορούν το κομμάτι του δικτύου αισθητήρων και β) λειτουργίες που εκτελούνται στο κέντρο τηλεελέγχου ή στο cloud. Η ανάλυση των δεδομένων υλοποιείται σε δύο σημεία: στα έξυπνα υδρόμετρα (τα υδρόμετρα που έχει προμηθευτεί η ΔΕΥΑΜ υποστηρίζουν εγγενώς λειτουργίες edge processing, βλ. Παραδοτέο 3.3) και στα κέντρα τηλεελέγχου των αντίστοιχων εταίρων του έργου (ΣΥΛΑ, ΣΥΛΕ, ΤΑΥ, ΔΕΥΑΜ). Η ανακατασκευή των δεδομένων, η παρακολούθηση συσχετίσεων μεταξύ ροών δεδομένων και η αναγνώριση μη φυσιολογικής συμπεριφοράς υλοποιούνται στα κέντρα τηλεελέγχου. Η συμπίεση των δεδομένων εφαρμόζεται στα έξυπνα υδρόμετρα, τα οποία διαχειρίζονται την ολότητα των δεδομένων.

### Δίκτυο αισθητήρων

- Συμπιεστική δειγματοληψία
- Έλεγχος βαλβίδων πίεσης
- Ανίχνευση διαρροών
- Ανίχνευση ανωμαλιών στη συγκέντρωση χλωρίου
- Έλεγχος πίεσης δικτύου

### Κέντρο τηλεελέγχου

- Ανακατασκευή δεδομένων
- Παρακολούθηση συσχετίσεων
- Αύξηση χρονικής ανάλυσης των λαμβανόμενων ροών δεδομένων
- Ανίχνευση μη φυσιολογικής συμπεριφοράς
- Παραγωγή ειδοποιήσεων
- Εκτίμηση Υδραυλικής κατάστασης δικτύου
- Αξιολόγηση ρίσκου σπασιμάτων σωληνών
- Εκτίμηση ποιοτικής κατάστασης του νερού του δικτύου

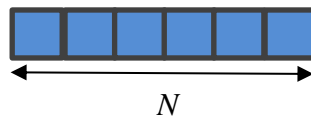
### 3 Αλγόριθμοι ανάλυση σήματος από δίκτυα αισθητήρων

Στην ενότητα αυτή περιγράφονται αναλυτικά οι επιμέρους λειτουργίες και οι αντίστοιχοι αλγόριθμοι. Πιο συγκεκριμένα, δίνεται μια περιγραφή του επιλεγμένου αλγορίθμου, καθώς και των δεδομένων που χρησιμοποιούνται για τον έλεγχο της απόδοσής του (π.χ. input/output, πραγματικά ή προσομοιωμένα με LeakDB).

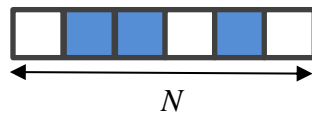
#### 3.1 Σημειογραφία

Οι παρακάτω συμβολισμοί και ορισμοί χρησιμοποιούνται στη συνέχεια για την αναλυτική περιγραφή των παραπάνω λειτουργιών.

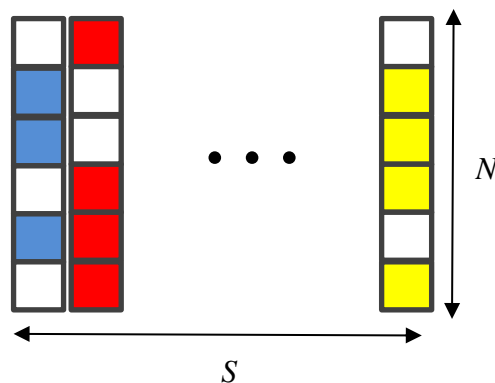
Ως διάνυσμα μετρήσεων ορίζουμε μια ακολουθία τιμών που έχουν ληφθεί από έναν αισθητήρα σε μια χρονική περίοδο  $T$ . Με  $N$  συμβολίζεται το πλήθος των μετρήσεων αυτών.



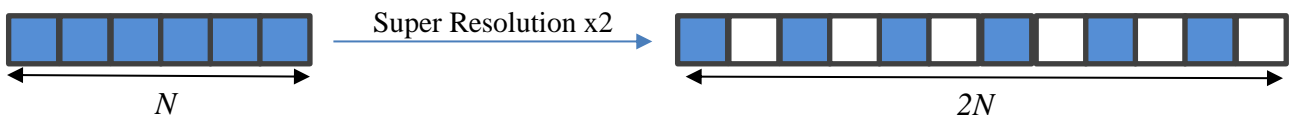
Τα δεδομένα των αισθητήρων εξετάζονται σε χρονικά παράθυρα μήκους  $w$ . Στο διάστημα αυτό, λόγω αδυναμίας του δικτύου ή περιορισμένων αποθεμάτων μπαταρίας, ενδέχεται να υπάρχουν μη καταγεγραμμένες μετρήσεις. Ως εκ τούτου, το διάνυσμα δεδομένων που λαμβάνεται από το κέντρο τηλεελέγχου μπορεί να περιέχει χαμένες τιμές,



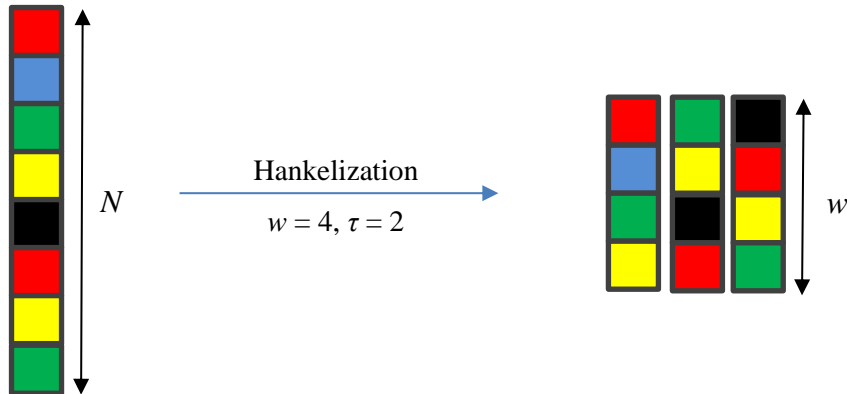
Οι μετρήσεις όλων των αισθητήρων λαμβάνονται από το κέντρο τηλεελέγχου και αποθηκεύονται σε διδιάστατες δομές, τους πίνακες μετρήσεων. Έτσι, κάθε στήλη αναπαριστά τα δεδομένα καθενός από τους  $S$  αισθητήρες ενώ οι γραμμές τις  $N$  μετρήσεις των αισθητήρων. Ο πίνακας μετρήσεων ενδέχεται να περιλαμβάνει μη καταγεγραμμένες μετρήσεις.



Η αύξηση της χρονικής ανάλυσης έχει ως στόχο την εξέταση των δεδομένων σε μια συχνότητα διαφορετική από τη συχνότητα δειγματοληψίας. Μπορούμε δηλαδή να εξετάσουμε τα δεδομένα σε κάποιο πολλαπλάσιο της συχνότητας με την οποία λαμβάνονται. Αν για παράδειγμα η δειγματοληψία πραγματοποιείται κάθε 30 λεπτά και επιθυμούμε να τα εξετάσουμε σε συχνότητα 15 λεπτών, τότε δημιουργείται ένα ενδιάμεσο διάνυσμα όπως φαίνεται στο παρακάτω σχήμα,



Ένας ακόμα τελεστής που χρησιμοποιείται σε διανύσματα δεδομένων είναι η μετατροπή τους σε πίνακα Hankel (Hankelization). Πιο συγκεκριμένα, δοθέντος διανύσματος μήκους  $N$ , παραθύρου μεγέθους  $w$  και ενός βήματος  $\tau$ , προκύπτει πίνακας Hankel όπως φαίνεται παρακάτω. Παράλληλα, υπάρχει η δυνατότητα μετατροπής στο αρχικό διάνυσμα, μέσω της αντίστροφης διαδικασίας (Dehankelization).



## 3.2 Συμπιεστική δειγματοληψία

Η συμπιεστική δειγματοληψία πραγματοποιείται σε δύο επίπεδα: στα άκρα του δικτύου και στο κέντρο τηλεελέγχου. Αρχικά γίνεται η συμπίεση των δεδομένων με σκοπό τη μείωση του όγκου τους πριν την αποστολή στο κέντρο τηλεελέγχου. Ο μειωμένος όγκος δεδομένων οδηγεί στη μείωση του κόστους τηλεμετρίας, καθώς και στην αύξηση της αυτονομίας του συστήματος. Αφού ολοκληρωθεί η λήψη των συμπιεσμένων δεδομένων στο κέντρο τηλεελέγχου, τα δεδομένα αποσυμπιέζονται προκειμένου να επεξεργαστούν.

### 3.2.1 Συμπίεση στα άκρα του δικτύου

Η συμπίεση των δεδομένων υλοποιείται στα άκρα του δικτύου με σκοπό τη μείωση του όγκου τους. Έστω  $\mathbf{x}$  το αρχικό διάνυσμα δεδομένων μήκους  $N$ . Το συμπιεσμένο διάνυσμα  $\mathbf{y}$  μήκους  $M$  δίνεται από τη σχέση  $\mathbf{y} = \Phi \mathbf{x}$ . Ο πίνακας  $\Phi$ , διαστάσεων  $M \times N$  μοντελοποιεί τη διαδικασία γραμμικής υποδειγματοληψίας και ορίζεται βάσει των παραμέτρων seed και PhiType. Το PhiType ορίζει τον τύπο του πίνακα  $\Phi$  (π.χ. Gaussian, Bernouli), ενώ το seed αποτελεί την πηγή παραγωγής της γεννήτριας ψευδοτυχαίων αριθμών που συνιστούν τα στοιχεία του πίνακα.

**Function:** `CSEdge(x, seed, PhiType, samplingRatio)`

#### Inputs:

@x: το αρχικό διάνυσμα με όλες τις μετρήσεις των αισθητήρων μήκους  $N$

@seed: ακέραιος που καθορίζει την ακολουθία που παράγει η γεννήτρια ψευδοτυχαίων αριθμών για τη δημιουργία του πίνακα  $\Phi$

@PhiType: ο τύπος του πίνακα δειγματοληψίας  $\Phi$

@samplingRatio: το ποσοστό δειγματοληψίας των δεδομένων. Το μήκος του συμπιεσμένου διανύσματος υπολογίζεται ως  $M = \text{floor}(\text{samplingRatio} * N)$

#### Output:

@y: το συμπιεσμένο διάνυσμα δεδομένων μήκους  $M$  ( $M \ll N$ )

Variable	Type	Default Value	Description
$\mathbf{x}$	ArrayList<Double>	-	Input data

<b>seed</b>	int		Sampling matrix seed
<b>PhiType</b>	string	“rsenn” (Bernouli)	Sampling matrix type
<b>samplingRatio</b>	double	0.6	Sampling ratio in (0.0,1.0]
<b>y</b>	ArrayList<Double>	-	Compressed data stream

### 3.2.2 Αποσυμπίεση στο κέντρο τηλεελέγχου

Στο κέντρο τηλεελέγχου πραγματοποιείται η αποσυμπίεση (ανακατασκευή) των δεδομένων προκειμένου να γίνει περαιτέρω επεξεργασία τους. Για την αποσυμπίεση των δεδομένων λύνεται κατάλληλο πρόβλημα βελτιστοποίησης με περιορισμούς (βλ. Παραδοτέο 3.2). Ο τελεστής αραιοποίησης που χρησιμοποιούμε στην παρούσα υλοποίηση είναι ο Short-Time Fourier Transform (STFT), ενώ για την ανακατασκευή βασίζομαστε στον αλγόριθμο NESTA<sup>1</sup>.

**Function:** *CSDecompression*(y, seed, PhiType, psi)

#### Inputs:

@y: το συμπίεσμένο διάνυσμα δεδομένων μήκους  $M$

@seed: ακέραιος που καθορίζει την ακολουθία που παράγει η γεννήτρια ψευδοτυχαίων αριθμών για τη δημιουργία του πίνακα  $\Phi$

@PhiType: ο τύπος του πίνακα δειγματοληψίας  $\Phi$

@psi: δομή για τη δημιουργία του τελεστή αραιοποίησης, η οποία περιλαμβάνει

@psi.gMax: ακέραιος που καθορίζει τη μέγιστη κλίμακα του STFT (max scale =  $\log N + gMax$ )

@psi.gLevels: ακέραιος που καθορίζει την ελάχιστη κλίμακα του STFT (min scale =  $\log N - gLevels$ )

@psi.tRedundancy: ακέραιος που καθορίζει τη μετατόπιση του STFT παραθύρου στο χρόνο

@psi.fRedundancy: ακέραιος που καθορίζει τη μετατόπιση του STFT παραθύρου στη συχνότητα

@psi.gWindow: τύπος παραθύρου (η παρούσα υλοποίηση υποστηρίζει την επιλογή ‘iterate sine’)

#### Output:

@x: το αποσυμπίεσμένο διάνυσμα δεδομένων μήκους  $N$

Variable	Type	Default Value	Description
<b>y</b>	HashMap<String,ArrayList<Double>>	-	Input data
<b>seed</b>			Sampling matrix seed
<b>PhiType</b>	string	“rsenn” (Bernouli)	Sampling matrix type
<b>psi</b>	psiStruct		

<sup>1</sup> <https://statweb.stanford.edu/~candes/nesta>

<b>psi.gMax</b>	int	0	Determines max scale of STFT (= $\log_2 N + gMax$ )
<b>psi.gLevel</b>	int	$\text{floor}(\log_2(N)/2)$	Determines min scale of STFT (= $\log_2 N - gLevels$ )
<b>psi.tRedundancy</b>	int	2	Shift of window in time
<b>psi.fRedundancy</b>	int	1	Shift of window in frequency
<b>psi.gWindow</b>	string	'isine'	Window type

### 3.3 Παρακολούθηση συσχετίσεων

Οι συσχετίσεις εκφράζουν τη στατιστική σχέση μεταξύ δύο ροών δεδομένων. Στη λειτουργία αυτή υπολογίζονται οι συσχετίσεις μεταξύ όλων των δυνατών ζευγών. Όταν η τιμή της συσχέτισης υπερβαίνει κάποιο προκαθορισμένο όριο, τότε οι αντίστοιχες ροές δεδομένων χαρακτηρίζονται ως υψηλά συσχετισμένες. Για την επιτάχυνση της εκτέλεσης της συνάρτησης, σε περίπτωση που το πλήθος των δεδομένων  $N$  είναι μεγάλο, ο υπολογισμός της συσχέτισης πραγματοποιείται στο πεδίο της συχνότητας μετασχηματίζοντας τις λαμβανόμενες ροές μέσω του truncated Fourier transform. Η περικοπή (truncation) των πολύ υψηλών συχνοτήτων έχει ως αποτέλεσμα την ελαφρά μείωση της ακρίβειας στον υπολογισμό της συσχέτισης. Για την αντιμετώπιση αυτού του προβλήματος εισάγεται ένας συντελεστής διόρθωσης στη τιμή της συσχέτισης (βλ. Παραδοτέο 3.2). Η παρακολούθηση των συσχετίσεων ενισχύει την έγκαιρη και έγκυρη ανίχνευση της μη φυσιολογικής συμπεριφοράς και παραγωγής ειδοποιήσεων, καθώς επίσης χρησιμοποιείται σαν βήμα προεπεξεργασίας των ροών πριν την εκτέλεση του αλγορίθμου για την ανάκτηση χαμένων τιμών ή την αύξηση της χρονικής ανάλυσης (ο αλγόριθμος συμπλήρωσης τανυστή εκτελείται πάνω σε συσχετισμένες ροές δεδομένων).

**Function:** `CorrelationMonitoring(X, threshold, thresholdCorrection, fourierCutoff, energyCutoff)`

#### Inputs:

@X: πίνακας με στήλες τα διανύσματα με τις μετρήσεις των αισθητήρων

@threshold: όριο πάνω από το οποίο δύο ροές δεδομένων χαρακτηρίζονται ως υψηλά συσχετισμένες

@thresholdCorrection: τιμή που προστίθεται στην υπολογισμένη συσχέτιση (στο πεδίο της συχνότητας) για τη διόρθωση της συσχέτισης στο πεδίο του χρόνου

@fourierCutoff: τιμή που χρησιμοποιείται για τον προσδιορισμό του μήκους των διανυσμάτων, στο πεδίο της συχνότητας, που θα χρησιμοποιηθούν για τον υπολογισμό των συσχετίσεων (π.χ. 2 για περικοπή στο  $N/2$  των συχνοτήτων, 4 για  $N/4$ , κλπ.)

@energyCutoff: ποσοστό της ενέργειας στο πεδίο των συχνοτήτων που καθορίζει το πλήθος των δεδομένων που θα χρησιμοποιηθούν για τον υπολογισμό της συσχέτισης

#### Output:

@Y: πίνακας με τις υπολογισμένες συσχετίσεις μεταξύ όλων των ζευγών ροών δεδομένων



Variable	Type	Default Value	Description
<b>X</b>	HashMap<String,ArrayList<Double>>	-	Input data matrix
<b>threshold</b>	double	0.8	Highly correlated streams threshold
<b>thresholdCorrection</b>	double	0.065	Threshold correction
<b>fourierCutoff</b>	int	2	New stream size in the frequency domain ( $N/\text{fourierCutoff}$ )
<b>energyCutoff</b>	double	0.35	Sampling ratio in the frequency domain (0.0,1.0]
<b>Y</b>	HashMap<String,ArrayList<Double>>	-	Pairwise correlation map

### 3.4 Ανακατασκευή χαμένων μετρήσεων

#### 3.4.1 Έλεγχος χαμένων μετρήσεων

Η λειτουργία αυτή είναι υπεύθυνη για τον έλεγχο ύπαρξης χαμένων μετρήσεων με σκοπό την ανακατασκευή τους με χαμηλό υπολογιστικό κόστος. Οι χαμένες μετρήσεις αναπαρίστανται με NaN για την ομαλή λειτουργία των συναρτήσεων συμπλήρωσης των χαμένων μετρήσεων, η οποία επιτυγχάνεται με χρήση παρεμβολής με τμηματικά κυβικά πολυώνυμα Hermite.

**Function:** `checkMissingData(x, missingID)`

#### Inputs:

@x: το αρχικό διάνυσμα μετρήσεων

@missingID: συμβολοσειρά που καθορίζει τον τρόπο αναπαράστασης των χαμένων μετρήσεων (π.χ. NaN, NA)

#### Output:

@y: το διάνυσμα μετρήσεων με τις χαμένες μετρήσεις να έχουν συμπληρωθεί

Variable	Type	Default Value	Description
<b>x</b>	ArrayList<Double>	-	Input data stream
<b>missingID</b>	string	NaN	Missing value representation
<b>y</b>	ArrayList<Double>	-	Output stream without missing values

#### 3.4.2 Συμπλήρωση χαμένων μετρήσεων μεταξύ υψηλά συσχετισμένων ροών

Για ροές δεδομένων με υψηλή συσχέτιση μπορούν να εφαρμοστούν πιο εξελιγμένοι αλγόριθμοι συμπλήρωσης τιμών. Συγκεκριμένα, οι μετρήσεις μπορούν να αποθηκευτούν σε τριδιάστατες δομές, όπως οι τανυστές  $3^{\text{ης}}$  τάξης, και να εφαρμοστεί η μέθοδος συμπλήρωσης τανυστών. Ο αλγόριθμος που υλοποιήσαμε για τη συμπλήρωση τανυστών βασίζεται στον TMac<sup>2</sup>. Εφαρμόζοντας τον αλγόριθμό μας σε πραγματικά

<sup>2</sup> <https://xu-yangyang.github.io/TMac>

δεδομένα πίεσης που παραχώρησε η ΔΕΥΑΜ, παρατηρήσαμε ότι, στην πλειοψηφία των περιπτώσεων, η ακρίβεια της ανακατασκευής αυξάνεται όταν εφαρμόζεται ένα επιπλέον βήμα εξομάλυνσης μέσω μιας μεθόδου ελαχίστων τετραγώνων με ποινή<sup>3</sup>.

**Function:** **TCRecovery**(*X*, *windowSize*, *stepSize*)

**Inputs:**

@*X*: πίνακας με στήλες τις ροές δεδομένων από υψηλά συσχετισμένους αισθητήρες

@*windowSize*: μέγεθος παραθύρου που χρησιμοποιείται για το Hankelization των δεδομένων

@*stepSize*: βήμα που χρησιμοποιείται για τη μετάβαση στο επόμενο παράθυρο των Hankelized δεδομένων

**Output:**

@*Y*: πίνακας με στήλες τις ροές με συμπληρωμένες τις χαμένες μετρήσεις

Variable	Type	Default Value	Description
<b>X</b>	HashMap<String,ArrayList<Double>>	-	Input data matrix
<b>windowSize</b>	int	-	Hankelization window size
<b>stepSize</b>	int	-	Hankelization step size
<b>Y</b>	HashMap<String,ArrayList<Double>>	-	Output matrix without missing values

### 3.5 Αύξηση της χρονικής ανάλυσης

Η συγκεκριμένη λειτουργία έχει ως στόχο την τεχνητή αύξηση της συχνότητας δειγματοληψίας των δεδομένων με σκοπό την αύξηση της χρονικής ανάλυσής τους. Η επεξεργασία των δεδομένων σε μεγαλύτερες χρονικές αναλύσεις μπορεί να οδηγήσει στην εξαγωγή ακριβέστερων συμπερασμάτων για την κατάσταση του συστήματος. Αν, για παράδειγμα, η δειγματοληψία πραγματοποιείται με περίοδο  $T$  και σε κάθε παράθυρο λαμβάνουμε  $N$  μετρήσεις, μια αύξηση στο χρόνο  $x2$  θα οδηγούσε σε φαινομενική μεταβολή της περιόδου σε  $T/2$  και πλήθος μετρήσεων  $2N$  για το ίδιο χρονικό παράθυρο. Οι ενδιάμεσες μη καταγεγραμμένες τιμές που προκύπτουν εισάγονται με δομημένο τρόπο στο διάνυσμα μετρήσεων. Στη συνέχεια, τα διανύσματα ανακατασκευάζονται με τις μεθόδους που περιγράφονται παραπάνω.

**Function:** **SuperResolution**(*x*, *delta*)

**Inputs:**

@*x*: αρχικό διάνυσμα μετρήσεων μήκους  $N$

@*delta*: ακέραιο πολλαπλάσιο της αρχικής συχνότητας δειγματοληψίας των μετρήσεων

**Output:**

@*y*: ανακατασκευασμένο διάνυσμα με  $delta \cdot N$  μετρήσεις

Variable	Type	Default Value	Description
<b>x</b>	HashMap<String,ArrayList<Double>>	-	Input data stream of size $N$

<sup>3</sup> D. Garcia, “Robust smoothing of gridded data in one and higher dimensions with missing values,” *Comput. Statist. Data Anal.*, 54:1167-1178, 2010.

<b>delta</b>	int		Zoom coefficient
<b>y</b>	HashMap<String,ArrayList<Double>>	-	Output reconstructed zoomed stream of size $\text{delta} \cdot N$

Alert ID	Description	
0	OK	Green
1	Δεν έχουμε δεδομένα	Red
2	Σταθερή τιμή σε όλο το παράθυρο (βλάβη αισθητήρα)	Red
3	Data $\pm$ U κάτω από το κάτω όριο	Red
4	Data $\pm$ U πάνω από το άνω όριο	Red
5	Data - U κάτω από το κάτω όριο & Data + U μεταξύ των ορίων	Yellow
6	Data + U πάνω από το άνω όριο & Data - U μεταξύ των ορίων	Yellow
7	Data $\pm$ U εκτός και των δύο ορίων	Red

### 3.6 Αναγνώριση μη φυσιολογικής συμπεριφοράς

Η λειτουργία αυτή στοχεύει στον εντοπισμό της μη φυσιολογικής συμπεριφοράς στα δεδομένα (π.χ. πίεσης ή ποιότητας νερού) και την άμεση παραγωγή ειδοποιήσεων προς τους διαχειριστές του συστήματος. Η αξιοπιστία των ειδοποιήσεων αυξάνεται όταν ληφθεί υπόψη η εγγενής αβεβαιότητα των δεδομένων ή/και των υπολογιστικών μοντέλων. Συγκεκριμένα, για τον υπολογισμό της χρονικά μεταβαλλόμενης αβεβαιότητας σε δεδομένα πίεσης χρησιμοποιούνται τα υδραυλικά μοντέλα που αναπτύσσει η ομάδα του ΚΟΙΟΣ. Κατόπιν, οι τιμές των μετρήσεων, επαυξημένες με την αντίστοιχη αβεβαιότητα, συγκρίνονται με σταθερά ή μεταβλητά όρια που σχετίζονται με προκαθορισμένες τιμές ορθής λειτουργίας του συστήματος. Έτσι, προκύπτουν 8 διαφορετικά είδη ειδοποιήσεων. Για λόγους ευκολίας, οι ειδοποιήσεις ομαδοποιούνται σε 3 κατηγορίες βάσει της αναγκαιότητας λήψης αποφάσεων. Η πρώτη κατηγορία αφορά τη φυσιολογική λειτουργία του συστήματος (πράσινο χρώμα), η δεύτερη αφορά την περίπτωση όπου τα δεδομένα, συμπεριλαμβανομένης της αβεβαιότητας, βρίσκονται εντός των ορίων (πορτοκαλί χρώμα), και η τρίτη κατηγορία την περίπτωση στην οποία τα δεδομένα, συμπεριλαμβανομένης της αβεβαιότητας, βρίσκονται εκτός των ορίων.

## 4 Αλγόριθμοι για ανάλυση δικτύου, έλεγχος και παρακολούθηση

Στην ενότητα αυτή περιγράφονται αναλυτικά οι επιμέρους λειτουργίες και οι αντίστοιχοι αλγόριθμοι. Πιο συγκεκριμένα, δίνεται μια περιγραφή του επιλεγμένου αλγορίθμου, καθώς και των δεδομένων που χρησιμοποιούνται για τον έλεγχο της απόδοσής του (π.χ. input/output, πραγματικά ή προσομοιωμένα με LeakDB).

### 4.1 Εκτίμηση υδραυλικής κατάστασης δικτύου

Ο εκτιμητής υδραυλικής κατάστασης δικτύου υπολογίζει την πλήρη κατάσταση του δικτύου (ροή σε όλες τις σωλήνες, πίεση σε όλους τους κόμβους, επίπεδο νερού στις δεξαμενές) χρησιμοποιώντας τις διαθέσιμες μετρήσεις και ένα μαθηματικό μοντέλο του δικτύου. Μια ολοκληρωμένη εικόνα του δικτύου υποστηρίζει την διαδικασία λήψης αποφάσεων, καθιστά δυνατή τη βελτιστοποίηση της λειτουργίας του συστήματος διανομής νερού και επιτρέπει την έγκαιρη ανίχνευση σφαλμάτων όπως είναι οι διαρροές.

Το πρόβλημα της εκτίμησης υδραυλικής κατάστασης ενός δικτύου νερού εκφράζεται ως ένα μη-γραμμικό πρόβλημα μαθηματικού προγραμματισμού ως εξής:

$$\begin{aligned} \min_{h_i} \quad & \sum_{\forall i} \left( \frac{\tilde{h}_i - h_i}{\sigma_i^h} \right) + \sum_{\forall ij} \left( \frac{\tilde{Q}_{ij} - Q_{ij}}{\sigma_{ij}^Q} \right) + \sum_{\forall i} \left( \frac{\tilde{q}_i - q_i}{\sigma_i^q} \right) \\ \text{s.t.} \quad & f_h(h, Q, q) = 0 \\ & h^{\min} \leq h \leq h^{\max} \\ & Q^{\min} \leq Q \leq Q^{\max} \\ & q^{\min} \leq q \leq q^{\max} \end{aligned}$$

Όπου  $h$ ,  $Q$  και  $q$  είναι οι εκτιμώμενες πιέσεις, ροές και καταναλώσεις αντίστοιχα,  $\tilde{h}$ ,  $\tilde{Q}$  και  $\tilde{q}$  είναι οι διαθέσιμες μετρήσεις των αντίστοιχων ποσοτήτων και  $f_h(h, Q, q)$  είναι οι υδραυλικές εξισώσεις του δικτύου που πρέπει να ικανοποιούνται. Οι υπερθέσεις  $\min$  και  $\max$  υποδηλώνουν γνωστά κατώτατα και ανώτατα όρια αντίστοιχα στις μεταβλητές κατάστασης.

**Function:** *HydSE*( $\tilde{x}$ , *model*, *timeInfo*, *sensorInfo*)

**Inputs:**

@  $\tilde{x}$ : διαθέσιμες χρονοσειρές μετρήσεων ( $h$ ,  $Q$ ,  $q$ )

@ *model*: μαθηματικό μοντέλο του δικτύου σε μορφή αρχείου .inp

@ *timeInfo*: χρονικές πληροφορίες χρονοσειρών

@ *sensorInfo*: προδιαγραφές αισθητήρων και αντιστοιχία με μαθηματικό μοντέλο

**Output:**

@  $x$ : υπολογισμένες χρονοσειρές κατάστασης ως εικονικοί αισθητήρες

@ *virtualsensorInfo*: Πληροφορίες εικονικών μετρήσεων και αντιστοιχία με μαθηματικό μοντέλο

### 4.2 Έλεγχος πίεσης δικτύου με χρονοδρομολόγηση βαλβίδων μείωσης πίεσης

Η χρονοδρομολόγηση βαλβίδων μείωσης πίεσης (Pressure Reduction Valves) είναι ο προγραμματισμός των ρυθμίσεων αυτών των βαλβίδων ούτως ώστε να αλλάζουν την παρεχόμενη πίεση σε κάθε μετρήσιμη υποπεριοχή (District Metered Area) ανάλογα με την ώρα τις ημέρας και τις διακυμάνσεις της κατανάλωσης

νερού. Η χρονοδρομολόγηση των ρυθμίσεων υπολογίζεται με κριτήριο να ικανοποιείται μια ελάχιστη τιμή πίεσης σε όλη την υποπεριοχή. Ο υπολογισμός των πιέσεων σε όλα τα σημεία γίνεται δυνατός με τη χρήση της εκτίμησης κατάστασης. Οι ρυθμίσεις θα παρέχονται κάθε 24 ώρες και θα έχουν διάρκεια 24 ωρών. Για τον υπολογισμό τους απαιτούνται τα δεδομένα εκτίμησης κατάστασης δικτύου των προηγούμενων 7 ημερών.

**Function:** *settingsPRV(x, timeInfo, sensorInfo, model, minPressure)*

**Inputs:**

@ *x*: χρονοσειρές κατάστασης δικτύου διάρκειας μιας εβδομάδας

@ *model* : μαθηματικό μοντέλο του δικτύου σε μορφή αρχείου .inp

@ *timeInfo*: χρονικές πληροφορίες χρονοσειρών

@ *sensorInfo*: προδιαγραφές αισθητήρων και αντιστοιχία με μαθηματικό μοντέλο

@ *minPressure*: ελάχιστη επιθυμητή πίεση σε όλα τα σημεία του δικτύου

**Output:**

@ *settings*: υπολογισμένες χρονοσειρές ρυθμίσεων βαλβίδων

### 4.3 Αξιολόγηση ρίσκου σπάσιμου σωλήνων

Η αξιολόγηση ρίσκου σπάσιμου σωλήνων υπολογίζει μια εκτίμηση των κύκλων κόπωσης κάθε αγωγού στο δίκτυο λαμβάνοντας υπόψη τις διακυμάνσεις της πίεσης και την ηλικία του. Το ρίσκο σπασίματος κάθε αγωγού δίνεται σε σχέση με τους υπόλοιπους αγωγούς του δικτύου. Η πληροφορία αυτή είναι χρήσιμη στους λειτουργούς υδατοπρομήθειας όταν μια διαρροή ανιχνευτεί στο δίκτυο.

Η βιώσιμη διαχείριση των δικτύων διανομής αστικών υδάτων θα πρέπει να περιλαμβάνει όχι μόνο νέες μεθόδους παρακολούθησης, επισκευής ή αντικατάστασης της γήρανσης της υποδομής, αλλά και (και κυρίως) εκτεταμένες μεθόδους για τη μοντελοποίηση των επιδεινούμενων συνθηκών υποδομής, την αξιολόγηση των ιστορικών δεδομένων περιστατικών και τον εγγενή κίνδυνο αποτυχίας, στρατηγικές επισκευής, απαρίθμηση του κόστους κύκλου ζωής και οπτική παρουσίαση των διαμορφωμένων συνθηκών μέσω γεωχωρικών βάσεων δεδομένων και συστημάτων γεωγραφικών πληροφοριών για την καλύτερη διαβίβαση των υποκείμενων μαθηματικών μοντέλων και μετρήσεων αξιολόγησης σε όλους τους συμμετέχοντες οργανισμούς και τους ιδιοκτήτες υποδομών. Οι οργανισμοί ύδρευσης θα πρέπει να βρίσκονται στο επίκεντρο κάθε αποτελεσματικής και βιώσιμης προσπάθειας για καλύτερη διαχείριση των δικτύων σωληνώσεων, δεδομένου ότι η επιτυχής επίλυση του προβλήματος θα συμβάλει στη μείωση των απωλειών ύδατος και του κόστους κύκλου ζωής καθώς και στην αύξηση της αξιοπιστίας του δικτύου και της ποιότητας των υπηρεσιών οι πολίτες.

Η έρευνα που παρουσιάστηκε στα Kanakoudis & Tsitsifli (2007, 2011), Tsitsifli & Kanakoudis (2010) and Tsitsifli et al. (2011) στοχεύει στην ανάπτυξη ενός μοντέλου που θα μπορούσε να ταξινομήσει σωστά τους σωλήνες διανομής νερού και στον ορισμό των χαρακτηριστικών των αγωγών. Ένα μοντέλο που στοχεύει συγκεκριμένα τμήματα δικτύου, που παρουσιάζονται από τους Jun et al. (2008), εντοπίζει σωλήνες και βαλβίδες μεγάλης σημασίας. Το μοντέλο βασίζεται σε έναν αλγόριθμο μήτρας που μπορεί εύκολα να εφαρμοστεί σε οποιοδήποτε δίκτυο νερού και σε αναλύσεις που μπορούν να εκτελεστούν αποτελεσματικά ανεξάρτητα από το μέγεθος του συστήματος διανομής νερού.

Στο Bentes et al. (2011), πρότειναν ένα μοντέλο στήριξης αποφάσεων που στοχεύει το επίπεδο κατάστασης των δικτύων νερού και μια μέθοδο για την αξιολόγηση της ευπάθειας τους. Η πρότασή τους βασίζεται σε τρία παραδείγματα δικτύου σωλήνων ύδρευσης, εστιάζοντας στον εντοπισμό των ασθενέστερων τμημάτων του συστήματος διανομής και παρέχοντας καθοδήγηση για τη βελτίωση της κατάστασης του δικτύου. Στο Xu et al (2011) παρουσιάστηκε ένα μαθηματικό μοντέλο που χρησιμοποιεί δύο τεχνικές που βασίζονται σε δεδομένα. τον γενετικό προγραμματισμό και την εξελικτική πολυωνυμική παλινδρόμηση. Το μοντέλο

χρησιμοποιεί δεδομένα σχετικά με τα καταγεγραμμένα περιστατικά διακοπής αγωγών για τους αγωγούς του συγκεκριμένου δικτύου. Προχωρώντας ένα βήμα μπροστά, ο Xu et al. (2013) παρουσίασε την ανάπτυξη ενός Decision Support System που χρησιμοποιείται για την ανάπτυξη οικονομικά αποδοτικών σχεδίων συντήρησης σωλήνων για το δίκτυο. Το μοντέλο περιλαμβάνει ένα εργαλείο πρόβλεψης το οποίο έχει αναπτυχθεί χρησιμοποιώντας γενετικό προγραμματισμό και ένα οικονομικά βέλτιστο εργαλείο αντικατάστασης σωλήνων. Μια ερευνητική έκθεση που δημοσιεύτηκε από τους Grigg et al. (2013) του Ιδρύματος Υδατικών Ερευνών (WRF) εξέτασε την ανάγκη αποκατάστασης των δικτύων νερού και παρουσίασε ένα εργαλείο για την αξιολόγηση του κινδύνου και την οργάνωση δεδομένων για την ενίσχυση του προγραμματισμού των επενδύσεων κεφαλαίου για τους αγωγούς.

Οι Muranho et al. (2014) παρουσίασαν ένα μοντέλο που εντοπίζει προβληματικές ζώνες εντός ενός δικτύου με τη χρήση τεχνικών δεικτών επιδόσεων (TPI). Το μοντέλο, το οποίο περιλαμβάνει εργαλεία ανάλυσης και μοντέλο προσομοίωσης υπό πίεση, μπορεί να χρησιμοποιηθεί για την αξιολόγηση απόδοσης ενός δικτύου κατά τη διάρκεια εκρήξεων σωλήνων ή σεναρίων πυρόσβεσης.

Οι Christodoulou et al. (2009) εργάστηκαν σε τεχνικές μοντελοποίησης με γνώμονα τα δεδομένα, όπως τα τεχνητά νευρωνικά δίκτυα (ANN) και τα νευροσυζυγικά συστήματα. Αυτές οι τεχνικές μοντελοποίησης δεδομένων επιτρέπουν την ενσωμάτωση πολλαπλών παραγόντων κινδύνου στην ανάλυση, την καθιστούν πιο περίπλοκη, αλλά ταυτόχρονα πιο ακριβή και ρεαλιστική. Προχωρώντας ένα βήμα μπροστά, ο Agathokleous (2015) εκμεταλλεύτηκε αυτά τα ευρήματα για την ανάπτυξη ενός μαθηματικού μοντέλου ανάλυσης επιβίωσης που προσομοιώνει τη συμπεριφορά του. Τα χρησιμοποιούμενα δεδομένα συσχετίζονται με δεδομένα αποτυχίας των σωλήνων WM και HC. Το έργο των Christodoulou et al. (2012) παρουσίασε ένα μοντέλο χωροχρονικής ανάλυσης που μπορεί να χρησιμοποιηθεί ως Decision Support System για την αύξηση της αποδοτικότητας των στρατηγικών συντήρησης που σχετίζονται με το WDN. Το προτεινόμενο μοντέλο χρησιμοποιεί κλασικά στατιστικά εργαλεία, συστήματα νευροφυσίων και τεχνικές ομαδοχρονικής ομαδοποίησης και οπτικοποίησης με βάση το GIS.

Η εργασία που χρησιμοποιεί δεδομένα σχετικά με τους σωλήνες HC και τα εξαρτήματα συναρμολόγησης του δικτύου, εκτός από τα παραδοσιακά χρησιμοποιούμενα δεδομένα (σωλήνες WM), παρουσιάστηκε από τους Agathokleous (2016a) και Agathokleous & Christodoulou (2016b). Το κύριο μέρος αυτής της εργασίας επικεντρώνεται στη μοντελοποίηση της συμπεριφοράς του WDN, συγκρίνοντας διάφορα μοντέλα που συνδυάζουν διαφορετικά επίπεδα δεδομένων.

#### 4.4 Ανίχνευση διαρροών

Η ανίχνευση διαρροών σε πραγματικό χρόνο με τη χρήση δεδομένων τηλεμετρίας μπορεί να γίνει με διάφορους τρόπους και αναλόγως των διαθέσιμων μετρήσεων.

Με τη χρήση αισθητήρων ροής στην είσοδο των μετρήσιμων υποπεριοχών, μπορούν να ανιχνευτούν ανωμαλίες στην ροή, ιδιαίτερα τις νυχτερινές ώρες κατά τις οποίες η κατανάλωση είναι πιο προβλέψιμη. Οι ανωμαλίες αυτές υποδεικνύουν πιθανό σφάλμα σε αυτή την υποπεριοχή.

Το πρώτο βήμα για την χρησιμοποίηση μετρήσεων πίεσης για ανίχνευση διαρροών είναι η βέλτιστη τοποθέτηση αισθητήρων ούτως ώστε να μεγιστοποιείται η ανιχνευσιμότητα των διαρροών. Με τη χρήση αισθητήρων πίεσης και του μαθηματικού μοντέλου του δικτύου είναι εφικτή η ανίχνευση διαρροών βρίσκοντας ασυμφωνίες μεταξύ των μετρήσεων και της αναμενόμενης κατάστασης του δικτύου. Επιπρόσθετα, αν υπάρχουν διαθέσιμοι αρκετοί αισθητήρες, η πιθανή τοποθεσία της διαρροής μπορεί να εκτιμηθεί.

Οι Αυτόματοι Μετρητές παρέχουν μετρήσεις σε επίπεδο καταναλωτή και μπορούν να χρησιμοποιηθούν για να ανιχνευτούν ανωμαλίες στην κατανάλωση. Εγκαθίστανται σε μεγάλους καταναλωτές των οποίων η κατανάλωση μπορεί να επηρεάσει την εύρυθμη λειτουργία του δικτύου. Η ανίχνευση ανωμαλιών γίνεται με επεξεργασία χρονοσειρών.

#### 4.4.1 Ανίχνευση διαρροών με ανάλυση Ελάχιστης Νυχτερινής Ροής

Ο αλγόριθμος ανίχνευσης διαρροών με ανάλυση Ελάχιστης Νυχτερινής Ροής χρησιμοποιά τις μετρήσεις ροής στις εισόδους των Μετρήσιμων υποπεριοχών για μια προκαθορισμένη περίοδο μιας εβδομάδας. Από αυτές τις μετρήσεις υπολογίζεται η ελάχιστη νυχτερινή ροή κάθε μέρας. Η λογική πίσω από αυτή τη προσέγγιση, είναι ότι κατά τη διάρκεια των νυχτερινών ωρών, οι καταναλώσεις μεταβάλλονται λιγότερο και έτσι μειώνεται η αβεβαιότητα. Παράλληλα, η χαμηλή κατανάλωση έχει ως αποτέλεσμα υψηλές πιέσεις οι οποίες μεγεθύνουν την επίδραση οποιασδήποτε πιθανής διαρροής στο δίκτυο.

Η μεθοδολογία ανίχνευσης κανονικοποιά τις μετρήσεις για να αφαιρεθούν οι εποχικές αλλαγές. Τότε χρησιμοποιώντας τις κανονικοποιημένες μετρήσεις νυχτερινής ροής, υπολογίζεται ένα κατώφλι το οποίο λαμβάνει υπόψη την αβεβαιότητα στις μετρήσεις. Το κατώφλι αυτό χρησιμοποιείται για να αξιολογηθεί η νυχτερινή ροή της αμέσως επόμενης νύχτας.

##### **Inputs:**

@ *sensorTS*: χρονοσειρές κατάστασης δικτύου διάρκειας μιας εβδομάδας

@*sensorInfo*: προδιαγραφές αισθητήρων και αντιστοιχία με μαθηματικό μοντέλο

@*timeInfo*: χρονικές πληροφορίες χρονοσειρών

@*parametersMNF*: παράμετροι αλγορίθμου ανίχνευσης

##### **Output:**

@*warning*: προειδοποίηση σε περίπτωση ανίχνευσης ανωμαλιών στη συγκεκριμένη υποπεριοχή

#### 4.4.2 Τοποθέτηση αισθητήρων πίεσης

Η τοποθέτηση αισθητήρων πίεσης γίνεται με κριτήριο τη μέγιστη κάλυψη του δικτύου (ανιχνευσιμότητα) αλλά και την ικανότητα προσδιορισμού θέσης των διαρροών (απομόνωση).

**Function:** *PSPlace(model)*

##### **Inputs:**

@*model* : μαθηματικό μοντέλο του δικτύου σε μορφή αρχείου .inp

##### **Output:**

@*nodeIDs*: αναγνωριστικός αριθμός κόμβων στους οποίους θα τοποθετηθούν αισθητήρες

#### 4.4.3 Ανίχνευση διαρροών με τη χρήση αισθητήρων πίεσης

Ο αλγόριθμος χρησιμοποιεί τα δεδομένα εκτίμησης κατάστασης προηγούμενων ημερών για να προβλέψει την κατάσταση του δικτύου την παρούσα χρονική στιγμή, χρησιμοποιώντας το μοντέλο του δικτύου. Κατά την εκτίμηση προστίθεται αβεβαιότητα στις καταναλώσεις αλλά και στο μοντέλο, με αποτέλεσμα οι εκτιμώμενες πιέσεις να χαρακτηρίζονται από ένα κατώτατο και ανώτατο όριο.

Οι ανανεωμένες μετρήσεις πιέσεις λαμβάνονται από τους αισθητήρες και συγκρίνονται με τα εκτιμώμενα όρια για να υπολογιστεί το λεγόμενο υπολειπόμενο, το οποίο είναι ίσο με μηδέν όταν δεν υπάρχει παραβίαση των ορίων, ενώ ισούται με την παραβίαση όταν αυτή υπάρχει. Ο αλγόριθμος ανιχνεύει πιθανή διαρροή όταν υπάρχει οποιοδήποτε υπολειπόμενο το οποίο είναι μεγαλύτερο του μηδέν.

Αν υπάρξει ανίχνευση, ξεκινά η διαδικασία υπολογισμού της πιθανής τοποθεσίας της διαρροής με ανάλυση υπογραφής σφάλματος. Η υπογραφή σφάλματος είναι ο συγκεκριμένος συνδυασμός υπολοίπων που δημιουργεί η κάθε διαρροή. Με τη χρήση του πίνακα υπογραφών σφαλμάτων Fs ο οποίος είναι προϋπολογισμένος και περιέχει τις υπογραφές όλων των σφαλμάτων, επιλέγονται όλα τα σφάλματα τα οποία έχουν

την συγκεκριμένη υπογραφή και εξάγεται το σύνολο των πιθανών κόμβων στους οποίους μπορεί να έχει συμβεί η διαρροή.

**Function:** *pressureLD*( $\tilde{x}$ ,  $x$ , *timeInfo*, *sensorInfo*, *model*, *minPressure*)

**Inputs:**

@  $\tilde{x}$ : διαθέσιμες χρονοσειρές μετρήσεων ( $h$ ,  $Q$ ,  $q$ )

@  $x$ : χρονοσειρές εκτίμησης κατάστασης δικτύου διάρκειας μιας εβδομάδας

@*model*: μαθηματικό μοντέλο του δικτύου σε μορφή αρχείου .inp

@*timeInfo*: χρονικές πληροφορίες χρονοσειρών

@*sensorInfo*: προδιαγραφές αισθητήρων και αντιστοιχία με μαθηματικό μοντέλο

@*parametersPLD*: παράμετροι αλγορίθμου ανίχνευσης

**Output:**

@*warning*: προειδοποίηση σε περίπτωση ανίχνευσης ανωμαλιών στη συγκεκριμένη υποπεριοχή

@*nodeIDs*: αναγνωριστικός αριθμός κόμβων στους οποίους πιθανό να υπάρχει διαρροή

#### 4.4.4 Ανίχνευση διαρροών με χρήση Μετρήσεων Αυτόματων Μετρητών

Ο αλγόριθμος ανίχνευσης διαρροών με χρήση Μετρήσεων Αυτόματων Μετρητών, χρησιμοποιά μάθηση μηχανής για να μάθει τα χαρακτηριστικά του σήματος κάθε μέτρησης και να αναγνωρίσει ανωμαλίες στο σήμα. Αρχικά, λαμβάνονται μετρήσεις συγκεκριμένης περιόδου και το σήμα αυτό κανονικοποιείται για να αφαιρεθούν εποχικές παρεκκλίσεις. Τότε ένας αλγόριθμος προσαρμοστικής προσέγγισης μαθαίνει το άγνωστο, χρονικά μεταβαλλόμενο, περιοδικό σήμα με τη χρήση ακτινωτών συναρτήσεων βάσης των οποίων οι παράμετροι ανανεώνονται βάση νέων μετρήσεων. Η λογική ανίχνευσης σφάλματος χρησιμοποιά τον αλγόριθμο συσσωρευτικού αθροίσματος, ο οποίος ανιχνεύει αλλαγές στη μέση τιμή του σήματος.

**Function:** *AMRLD*( $\tilde{x}$ , *timeInfo*, *sensorInfo*)

**Inputs:**

@  $\tilde{x}$ : διαθέσιμες χρονοσειρές μετρήσεων ( $h$ ,  $Q$ ,  $q$ )

@*timeInfo*: χρονικές πληροφορίες χρονοσειρών

@*sensorInfo*: προδιαγραφές αισθητήρων και αντιστοιχία με μαθηματικό μοντέλο

@*parametersAMRLD*: παράμετροι αλγορίθμου ανίχνευσης

**Output:**

@*warning*: προειδοποίηση σε περίπτωση ανίχνευσης ανωμαλιών στη συγκεκριμένη κατανάλωση

#### 4.5 Εκτίμηση ποιοτικής κατάστασης του νερού στο δίκτυο

Η ποιοτική κατάσταση του νερού στο δίκτυο ορίζεται ως η συγκέντρωση σε αυτό διάφορων χημικών ουσιών. Εδώ μας ενδιαφέρει να εκτιμήσουμε σε όλα τα σημεία του δικτύου την συγκέντρωση του ελεύθερου χλωρίου στο νερό. Η αλγόριθμος εκτίμησης της ποιότητας γίνεται με τη χρήση αισθητήρων ελεύθερου χλωρίου τοποθετημένους μέσα στο δίκτυο και με τη χρήση δυναμικού μαθηματικού μοντέλου της ποιότητας του νερού. Για να υπολογιστεί η συγκέντρωση χλωρίου,



#### 4.5.1 Τοποθέτηση αισθητήρων χλωρίνης

Η τοποθέτηση αισθητήρων χλωρίνης είναι ένα πρόβλημα το οποίο επιλύεται offline με τη χρήση του εργαλείου S-PLACE. Το εργαλείο S-PLACE αποτελείται από ένα σύνολο μεθόδων για τη δημιουργία και προσομοίωση σεναρίων μόλυνσης του νερού και υδραυλικών σεναρίων, για υπολογισμό των επιπτώσεων ενός συμβάντος μόλυνσης του νερού. Το κόστος μόλυνσης ορίζεται ως ο όγκος μολυσμένου νερού που θα καταναλωθεί. Χρησιμοποιώντας το εργαλείο αυτό, είναι δυνατό να κατασκευαστούν σενάρια τα οποία περιλαμβάνουν την αβεβαιότητα η οποία μπορεί να υπάρχει στις παραμέτρους του προβλήματος. Οι παράμετροι που αλλάζουν είναι η πηγή και διάρκεια της μόλυνσης, καθώς και υδραυλικοί παράμετροι όπως η ζήτηση νερού αλλά και χαρακτηριστικά των σωλήνων του δικτύου.

Το πρόβλημα της βέλτιστης τοποθέτησης αισθητήρων ποιότητας, διατυπώνεται ως ένα πρόβλημα βελτιστοποίησης πολλαπλών στόχων ως εξής:

$$\begin{aligned} Y = \arg \min_{\chi \in \{0,1\}^{N_s}} \quad & \{F_0(\chi), F_1(\chi; \Omega), F_2(\chi, \Omega)\} \\ \text{s.t.} \quad & |\chi| \in X \end{aligned}$$

Όπου  $\chi_i = 1$  όταν ένας αισθητήρας θα τοποθετηθεί στον κόμβο  $i$ ,  $\chi_i = 0$  όταν δεν θα τοποθετηθεί αισθητήρας και  $X$  είναι το σύνολο των κόμβων όπου μπορεί να μπει αισθητήρας. Οι στόχοι της βελτιστοποίησης οι οποίοι χρειάζεται να ελαχιστοποιηθούν είναι:  $F_0$  ο αριθμός των αισθητήρων που θα εγκατασταθούν,  $F_1$  το εκτιμώμενο μέσο κόστος μόλυνσης (το οποίο ορίζεται ως ο όγκος μολυσμένου νερού που θα καταναλωθεί) και  $F_2$  το εκτιμώμενο χειρότερο κόστος μόλυνσης. Το πρόβλημα επιλύεται είτε με τη χρήση γενετικών αλγορίθμων είτε με απαρίθμηση και επίλυση όλων των πιθανών σεναρίων.

**Function:** *s\_place(model)*

**Inputs:**

@*model* : μαθηματικό μοντέλο του δικτύου σε μορφή αρχείου .inp

**Output:**

@*nodeIDs*: αναγνωριστικός αριθμός κόμβων στους οποίους θα τοποθετηθούν αισθητήρες

#### 4.5.2 Εκτίμηση συγκέντρωσης χλωρίου

Η εκτίμηση συγκέντρωσης χλωρίου μπορεί να χρησιμοποιηθεί για να αναγνωριστούν περιοχές στο δίκτυο όπου η απολύμανση του νερού δεν είναι ικανοποιητική. Η εκτίμηση της συγκέντρωσης χλωρίου γίνεται με τη χρήση προσομοίωσης της διασποράς και χημικών αντιδράσεων της χλωρίνης που εισέρχεται στο σύστημα. Για την εκτίμηση είναι αναγκαία η γνώση της ποσότητας χλωρίνης που εισέρχεται στο σύστημα και οι ροές στις σωλήνες για προκαθορισμένο διάστημα, καθώς και το ποιοτικό μοντέλο του δικτύου.

**Function:** *CISE(x̃, model, timeInfo, sensorInfo)*

**Inputs:**

@ *x̃* : διαθέσιμες χρονοσειρές εισόδου χλωρίνης και ροής νερού ( $u, Q$ )

@*model* : μαθηματικό μοντέλο του δικτύου σε μορφή αρχείου .inp

@*timeInfo*: χρονικές πληροφορίες χρονοσειρών

@*sensorInfo*: προδιαγραφές αισθητήρων και αντιστοιχία με μαθηματικό μοντέλο

**Output:**

@*x*: υπολογισμένες χρονοσειρές συγκέντρωσης χλωρίνης ως εικονικοί αισθητήρες

@*virtualsensorsInfo*: Πληροφορίες εικονικών μετρήσεων και αντιστοιχία με μαθηματικό μοντέλο

#### 4.6 Ανίχνευση ανωμαλιών στην συγκέντρωση χλωρίου

Η ανίχνευση ανωμαλιών στην συγκέντρωση χλωρίου γίνεται με τη χρήση κατάλληλα υπολογισμένων ορίων στην συγκέντρωση χλωρίνης στις περιοχές όπου υπάρχουν αισθητήρες χλωρίου. Όταν η μέτρηση από τους αισθητήρες ξεπεράσει τα υπολογισμένα όρια, είναι πιθανή ένδειξη μη αναμενόμενης ποιότητας νερού στο δίκτυο η οποία πιθανό να προέρχεται από κάποια μόλυνση ή μπορεί να σχετίζεται με κάποιο σφάλμα στη χλωρίωση ή στη λειτουργία των αισθητήρων.

Τα όρια συγκέντρωσης χλωρίνης υπολογίζονται λαμβάνοντας υπόψη την χλωρίνη που εισέρχεται στο σύστημα και τις υπολογισμένες ροές νερού, όπως και στον αλγόριθμο κατάστασης χλωρίνης. Εδώ όμως συνυπολογίζεται και η αβεβαιότητα στις ροές νερού καθώς και η αβεβαιότητα στο ποιοτικό μοντέλο του δικτύου, για να δοθεί ένα εύρος από αναμενόμενες τιμές συγκέντρωσης χλωρίου στους αισθητήρες το οποίο μπορεί να θεωρηθεί ως φυσιολογικό. Επίσης, διαφορετικά μεγέθη αβεβαιότητας, δίνουν όρια με πολλαπλά επίπεδα. Η παραβίαση κάθε επιπέδου συνοδεύεται με την ανάλογη πιθανότητα να συμβαίνει κάτι μη φυσιολογικό στην ποιότητα του νερού.

**Function:** *CIFD*( $\tilde{x}$ , *model*, *timeInfo*, *sensorInfo*)

##### **Inputs:**

@  $\tilde{x}$  : διαθέσιμες χρονοσειρές εισόδου χλωρίνης και ροής νερού ( $u$ ,  $Q$ )

@ *model* : μαθηματικό μοντέλο του δικτύου σε μορφή αρχείου .inp

@ *timeInfo*: χρονικές πληροφορίες χρονοσειρών

@ *sensorInfo*: προδιαγραφές αισθητήρων και αντιστοιχία με μαθηματικό μοντέλο

##### **Output:**

@  $x$ : υπολογισμένες χρονοσειρές συγκέντρωσης χλωρίνης ως εικονικοί αισθητήρες

@ *virtualsensorsInfo*: Πληροφορίες εικονικών μετρήσεων και αντιστοιχία με μαθηματικό μοντέλο

## 5 Σύνοψη

Το Παραδοτέο 3.4 αφορούσε τη περιγραφή βασικών λειτουργιών που έχουν αναπτυχθεί στις ανάγκες του έργου SmartWater2020. Οι βασικές λειτουργίες της πλατφόρμας του SmartWater2020 μπορούν να ταξινομηθούν σε δύο κατηγορίες: α) λειτουργίες που αφορούν το κομμάτι του δικτύου αισθητήρων και β) λειτουργίες που εκτελούνται στο κέντρο τηλεελέγχου. Η ανάλυση των δεδομένων υλοποιείται σε δύο σημεία: στα έξυπνα υδρόμετρα και στα κέντρα τηλεελέγχου των αντίστοιχων εταιρών του έργου. Η ανακατασκευή των δεδομένων, η παρακολούθηση συσχετίσεων μεταξύ ροών δεδομένων και η αναγνώριση μη φυσιολογικής συμπεριφοράς υλοποιούνται στα κέντρα τηλεελέγχου. Η συμπίεση των δεδομένων εφαρμόζεται στα έξυπνα υδρόμετρα, τα οποία διαχειρίζονται την ολότητα των δεδομένων. Από το ΙΤΕ και ΚΙΟΣ έχουν αναπτυχθεί Συμπιεστική δειγματοληψία, Έλεγχος βαλβίδων πίεσης, Ανίχνευση διαρροών μ Ανίχνευση ανωμαλιών στη συγκέντρωση χλωρίου, Έλεγχος πίεσης δικτύου, Ανακατασκευή δεδομένων, Παρακολούθηση συσχετίσεων, Αύξηση χρονικής ανάλυσης των λαμβανόμενων ροών δεδομένων, Ανίχνευση μη φυσιολογικής συμπεριφοράς, Παραγωγή ειδοποιήσεων, Εκτίμηση Υδραυλικής κατάστασης δικτύου, Αξιολόγηση ρίσκου σπασιμάτων σωληνών και Εκτίμηση ποιοτικής κατάστασης του νερού του δικτύου.

Όπως προβλέπει η συγκεκριμένη ερευνητική πρόταση, το επόμενο βήμα είναι η χρησιμοποίηση όλων των αναπτυγμένων βασικών λειτουργιών και με την χρήση των υδραυλικών μοντέλων από το Παραδοτέο 3.5 θα επιτευχθεί η εκτέλεση διάφορων υδραυλικών σεναρίων, ώστε να εντοπιστούν τα ακριβή σημεία στο δίκτυο όπου θα τοποθετηθούν αισθητήρες για βελτίωση και αναβάθμιση της παρακολούθησης του. Γενικά, τόσο οι βασικές λειτουργίες της πλατφόρμας όσο και τα υδραυλικά μοντέλα αποτελούν ουσιαστικό μέρος μιας ευρείας ποικιλίας συστημάτων υποστήριξης αποφάσεων για δίκτυα διανομής πόσιμου νερού. Αυτά τα μοντέλα μπορούν να ενσωματωθούν σε μια γενική πλατφόρμα η οποία είναι σε θέση να ενσωματώνει δεδομένα αισθητήρων σε πραγματικό χρόνο, και τις απαραίτητες πληροφορίες για την παραγωγή ρεαλιστικών εκτιμήσεων των ροών και πιέσεων του νερού μέσα στο δίκτυο.