

# ΑΝΑΛΥΣΗ ΑΦΙΞΕΩΝ ΠΡΟΣΦΥΓΙΚΩΝ ΡΟΩΝ ΓΙΑ ΕΞΥΠΙΝΑ ΣΥΣΤΗΜΑΤΑ ΕΓΚΑΙΡΗΣ ΠΡΟΕΙΔΟΠΟΙΗΣΗΣ ΤΩΝ ΥΠΗΡΕΣΙΩΝ ΥΠΟΣΤΗΡΙΞΗΣ ΣΤΑ ΝΗΣΙΑ ΠΡΩΤΗΣ ΥΠΟΔΟΧΗΣ

Χάρης Γεωργίου<sup>1</sup>, Ιωάννης Κιομουρτζής<sup>2</sup>, Φώτης Αλεξάκος<sup>3</sup>

<sup>1</sup>Γενικός Γραμματέας Δ.Σ. / <sup>2</sup>Αντιπρόεδρος Δ.Σ. / <sup>3</sup>Ειδικός Γραμματέας Δ.Σ., Ένωση Πληροφορικών Ελλάδας (Ε.Π.Ε.), Τ.Θ. 13801, Τ.Κ. 10310, Αθήνα

## Περίληψη

Η μελέτη παρουσιάζει για πρώτη φορά μια διεξοδική ανάλυση των καταγεγραμμένων ημερήσιων αφίξεων προσφυγικών ροών στην Ελλάδα για ένα διάστημα 15 εβδομάδων. Στόχος είναι η αναγνώριση προτύπων συμπεριφοράς στα παράνομα δίκτυα διακινητών, καθώς και των βασικών στατιστικών και περιοδικών χαρακτηριστικών των αφίξεων στα νησιά πρώτης υποδοχής. Διαπιστώνεται σαφέστατη περιοδική συμπεριφορά της τάξης των 6,2-6,5 ημερών, καθώς και ισχυρή εξάρτηση με τις 3 ως 4 προηγούμενες ημέρες (λιγότερο τις ενδιάμεσες). Η συμπεριφορά των δικτύων διακίνησης είναι συμβατή με αυτή των δικτύων δεδομένων «αποθήκευσης και προώθησης» (store-and-forward). Τέλος, αποδεικνύεται ότι η βραχυπρόθεσμη πρόβλεψη των αφίξεων είναι εν γένει εφικτή, με προοπτική ενσωμάτωσης ως έγκαιρη προειδοποίηση ή «επίπεδο επιφυλακής» σε συστήματα υποστήριξης των επιχειρήσεων έρευνας και διάσωσης στα σημεία πρώτης υποδοχής.

## Εισαγωγή

Η Ελλάδα αποτελεί την κύρια πύλη εισόδου των προσφυγικών ροών στην Ευρωπαϊκή Ένωση. Η έλλειψη συντονισμού και έγκαιρης προειδοποίησης ως προς τις αναμενόμενες αφίξεις συνέβαλε εμμέσως στα τραγικά αποτελέσματα πολύνεκρων ναυαγίων με χιλιάδες θύματα τους τελευταίους 18 μήνες.

Σύμφωνα με τις επίσημες καταγραφές του 2015, καθώς και για τους πρώτους μήνες του 2016, η αναλογία όσων διασώθηκαν από τη θάλασσα σε σχέση με το σύνολο των αφίξεων ήταν περισσότεροι από 1 στους 5. Από αυτούς, περίπου 1 στους 214 πνίγηκαν (ή αγνοούνται) στο Αιγαίο, τουλάχιστον 832 άνθρωποι συνολικά το 2015.

Στόχος της παρούσας μελέτης είναι να αναδείξει τη δυνατότητα εφαρμογής στατιστικής ανάλυσης και συστημικής μοντελοποίησης των ημερήσιων αφίξεων ως χρονοσειρά, έτσι ώστε να αναγνωριστούν οι βασικές τάσεις και περιοδικότητες του «σήματος». Επιπλέον, η κατασκευή μοντέλων για το σήμα ημερήσιων αφίξεων είναι δυνατό να αναδείξει τα ενδογενή χαρακτηριστικά των παράνομων δικτύων των διακινητών στις Τουρκικές ακτές, έτσι ώστε να προβλεφθεί η συμπεριφορά τους.

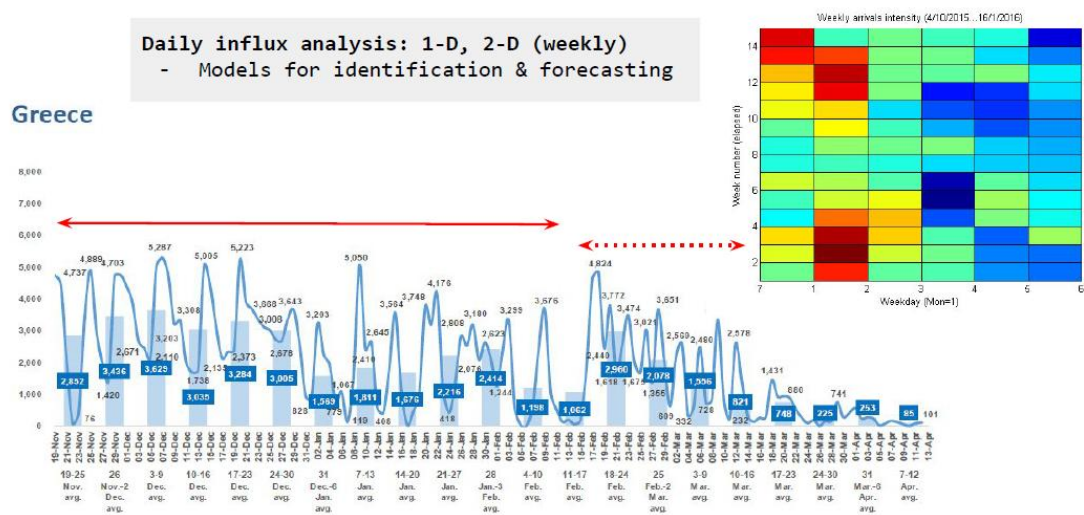
Τα χαρακτηριστικά αυτά, καθώς και οι αντίστοιχες μεθοδολογίες και αλγόριθμοι ανάλυσης, μπορούν να χρησιμοποιηθούν ως βάση για τη δημιουργία συστήματος έγκαιρης προειδοποίησης και «επίπεδου ετοιμότητας» για τις ομάδες διάσωσης που επιχειρούν στα νησιά πρώτης υποδοχής, με την προοπτική ενσωμάτωσής του σε αντίστοιχες πλατφόρμες λογισμικού που ήδη βρίσκονται σε επιχειρησιακή χρήση.

## Υλικά & Μέθοδοι

Ως βάση για την παρούσα μελέτη χρησιμοποιήθηκαν οι επίσημα καταγεγραμμένες ημερήσιες αφίξεις στα νησιά πρώτης υποδοχής (Λέσβος, Σάμος, Χίος, Κως, Λέρος, λοιπά Δωδεκάνησα), για μια περίοδο 15+ εβδομάδων, από αρχές Οκτωβρίου 2015 μέχρι μέσα Ιανουαρίου 2016 (1/10/2015-16/1/2016). Η περίοδος αυτή αποτελεί την πιο «θερμή» του 2015 από άποψη αφίξεων προσφυγικών ροών και φτάνει μέχρι πριν την αλλαγή της πολιτικής διαχείρισής τους, εν όψει της συμφωνίας ΕΕ-Τουρκίας που ακολούθησε λίγο αργότερα.

Οι μεθοδολογίες ανάλυσης περιλαμβάνουν, εκτός από τα βασικά στατιστικά κατανομής (Gaussian, Poisson, GEV), γραμμικές τάσεις (trends), περιοδικότητας, φασματική ανάλυση (DFT/FFT), γραμμική παλινδρόμηση (linear regression), γραμμική-συνημιτονοειδή παλινδρόμηση (cosine-linear regression), σύστημα αυτοσυσχέτισης (auto-correlation) και αυτοσυσχέτισης-εξομάλυνσης (ARMA). Επίσης τα δεδομένα αναλύονται σε διαφορετική δομή με ομαδοποίηση ανά εβδομάδα με αλγορίθμους παραγοντοποίησης πινάκων μέσω ιδιοτιμών (SVD), στατιστικής ανεξαρτησίας (ICA, Probabilistic PCA), καθώς και αραιών μοντέλων μέσω εκμάθησης λεξικού (K-SVD, Dictionary Learning). Τέλος, στο σήμα εφαρμόζεται χασοτική ανάλυση διάστασης (fractal dimension) για την εκτίμηση της στοχαστικότητας και της εν γένει δυνατότητας βραχυπρόθεσμης πρόβλεψης της εξέλιξής του (short-term forecasting).

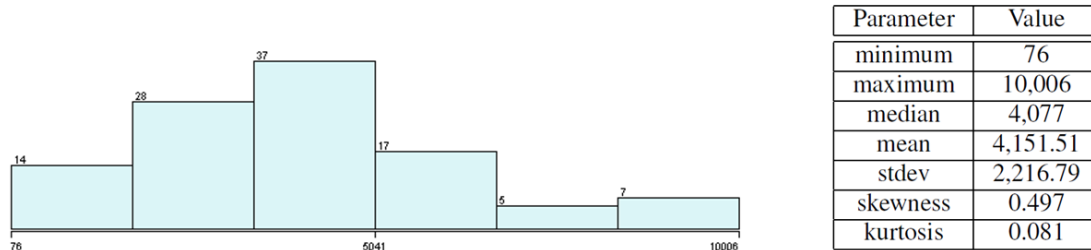
Στην Εικόνα 1 παρουσιάζεται το διάγραμμα των ημερήσιων αφίξεων, συμπεριλαμβάνοντας μέρος του διαστήματος που χρησιμοποιήθηκε στη μελέτη (Νοε/2015-Ιαν/2016, ενιαία γραμμή), το μεταβατικό διάστημα πριν τη συμφωνία μεταξύ ΕΕ-Τουρκίας (διακεκομμένη γραμμή) και το διάστημα αμέσως μετά από αυτή. Επίσης φαίνεται η εβδομαδιαία ομαδοποίηση του ρυθμού αφίξεων για το διάστημα της μελέτης, όπου διακρίνεται καθαρά η μεγαλύτερη ένταση το 48ωρο Κυριακής-Δευτέρας.



Εικόνα 1: Ημερήσιες αφίξεις και ομαδοποίηση ρυθμού αφίξεων ανά εβδομάδα.

## Αποτελέσματα

Στην Εικόνα 2 παρουσιάζονται τα βασικά στατιστικά (Gaussian moments) των ημερήσιων αφίξεων, καθώς και το ιστόγραμμα της πραγματικής κατανομής για το διάστημα της μελέτης (108 ημέρες).



Εικόνα 2: Ιστόγραμμα και βασικά στατιστικά (Gaussian) των ημερήσιων αφίξεων.

Σε ό,τι αφορά την ανάλυση μέσω παλινδρόμησης (regression), το γραμμικό μοντέλο έδειξε ισχυρή εξάρτηση με τις τέσσερις προηγούμενες ημέρες, ενώ το γραμμικό-συνημιτονοειδές μοντέλο φανέρωσε πως η τάση (trend) είναι -47 αφίξεις καθημερινά για το διάστημα που μελετήθηκε. Ο βασικός περιοδικός παράγοντας που εντοπίστηκε είναι 6,5 ημέρες, δηλαδή επαναλαμβανόμενο πρότυπο συμπεριφοράς σε σχεδόν εβδομαδιαία βάση, με μεγαλύτερη ένταση το 48ωρο Κυριακή-Δευτέρα.

$$\hat{y}(t) = (a \cdot \cos(b \cdot t + c)) + (d \cdot t + c_0)$$

$$\hat{y}(t) = (875 \cdot \cos(0.97 \cdot t - 2.85)) + (-47 \cdot t + 6669)$$

$$T_C = 2\pi/b \simeq \underline{6.5 \text{ (days)}}$$

### Cosine-linear Regression:

- Linear trend estim.
- Periodic trend estim.
- Major "frequency"
- High/Low peaks
- Very simple calc.

Εικόνα 3: Γραμμικό-συνημιτονοειδές μοντέλο παλινδρόμησης και οι βέλτιστες παράμετροι προσαρμογής (RMSE best-fit).

Σε ό,τι αφορά την ανάλυση αυτοσυσχέτισης (AR) και αυτοσυσχέτισης-εξομάλυνσης (ARMA), διαπιστώθηκε ισχυρή εξάρτηση με την τρίτη και τέταρτη προηγούμενη ημέρα, σε συμφωνία με τα προηγούμενα αποτελέσματα. Για την κατασκευή ενός αξιόπιστου μοντέλου βραχυπρόθεσμης πρόβλεψης, διαπιστώθηκε πως απαιτείται ένα συνεχές διάστημα περίπου τριών εβδομάδων (21 ημέρες) προηγούμενων σημείων. Η Εικόνα 4 παρουσιάζει τη βασική δομή ενός τέτοιου μοντέλου (ARMA) μικρότερης τάξης μεγέθους (n=9).

### ARMA modeling:

- Auto-regressive (y)
- Moving average (x)
- Sys. identification
- Short-term forecast
- Adaptive, simple

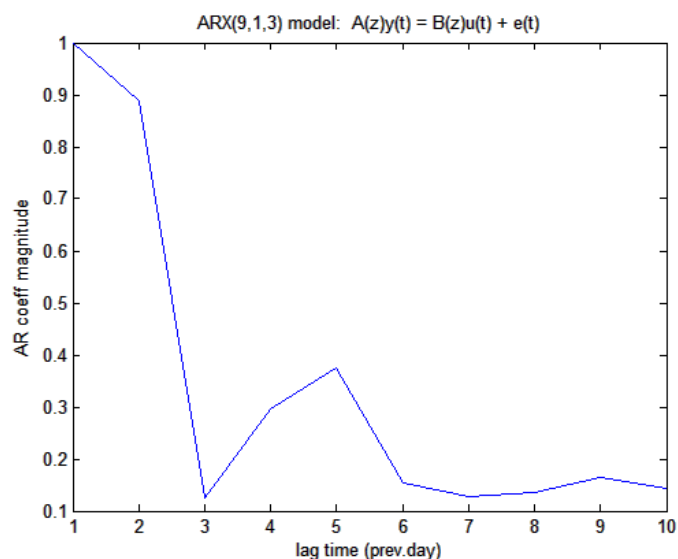
$$\hat{y}(t) = \sum_{i=1}^m (a_i \cdot y(t-i)) + \sum_{j=0}^k (b_j \cdot x(t-j)) + e(t)$$

$$A_9(z) = 1 - 0.8887 \cdot z^{-1} + 0.1247 \cdot z^{-2} + \underline{0.2971 \cdot z^{-3}} - 0.3747 \cdot z^{-4} + 0.1526 \cdot z^{-5} - 0.1265 \cdot z^{-6} - 0.1357 \cdot z^{-7} + 0.164 \cdot z^{-8} - 0.144 \cdot z^{-9}$$

$$B_9(z) = 48.94 \cdot z^{-3}$$

Εικόνα 4: Μοντέλο ARMA(9,1) και οι βέλτιστες παράμετροι (RMSE best-fit) για προσαρμογή στα δεδομένα ημερήσιων αφίξεων.

Για το παραπάνω μοντέλο, η Εικόνα 5 παρουσιάζει την κατανομή των συντελεστών αυτοσυσχέτισης, δηλαδή του AR(9), όπου είναι φανερό ότι η ισχυρότερη εξάρτηση εμφανίζεται στα σημεία -3 και -4 ημέρες από την τρέχουσα.

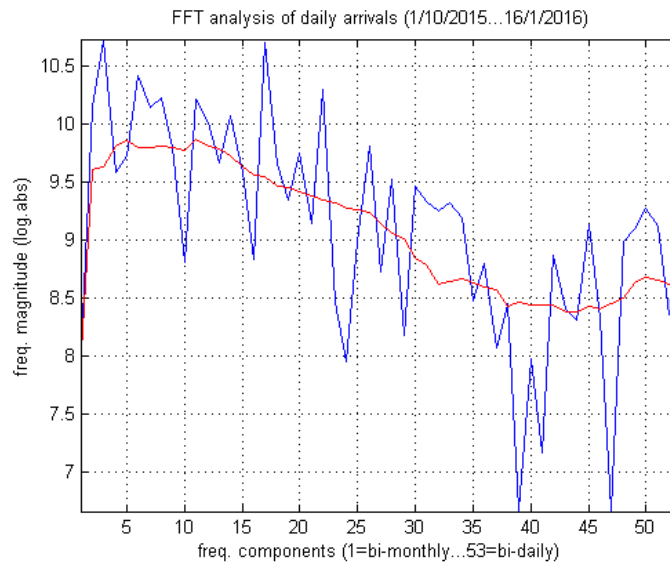


Εικόνα 5: Κατανομή των συντελεστών AR(9) για το βέλτιστο (RMSE best-fit) μοντέλο ARMA(9,1). Οι δείκτες στον οριζόντιο άξονα αντιστοιχούν σε ημέρες+1.

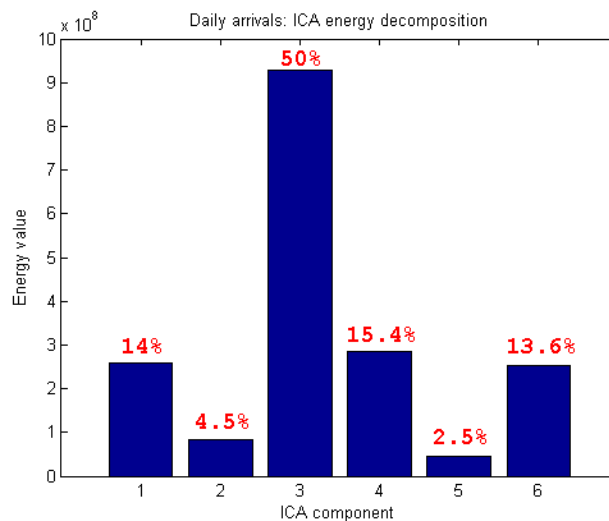
Η φασματική ανάλυση μέσω Fourier (DFT/FFT) ανέδειξε τις κύριες συνιστώσες του σήματος σε επίπεδο συχνοτήτων. Συγκεκριμένα, επιβεβαιώθηκε ότι οι ημερήσιες αφίξεις εμφανίζουν ισχυρή εξάρτηση από σχετικά χαμηλές συχνότητες, δηλαδή μικρή μεταβλητότητα μεταξύ συνεχόμενων ημερών. Εξετάζοντας τη φασματική απεικόνιση, όπως παρουσιάζεται στην Εικόνα 6, διαπιστώνεται μέσω υπολογισμών (αλλαγή κλίμακας στον οριζόντιο άξονα) ότι το σημείο που η φασματική πυκνότητα ενέργειας περιορίζεται οριστικά κάτω από το επίπεδο 9,5 αντιστοιχεί στις 6,2 ημέρες. Με άλλα λόγια, μέσω της φασματικής ανάλυσης επιβεβαιώνεται και πάλι ότι η βασική περιοδικότητα του σήματος βρίσκεται κάπου μεταξύ >6 και <7 ημέρες.

Σε ότι αφορά την εβδομαδιαία ομαδοποίηση και τους αλγορίθμους παραγοντοποίησης πινάκων, οι μέθοδοι SVD, PPCA και ICA, καθώς και τα αραιά μοντέλα μέσω K-SVD/DL, ανέδειξαν τα «φασματικά» πρότυπα συμπεριφοράς ανά εβδομάδα, δηλαδή τις βασικές τάσεις διάρκειας 7 ημερών που χαρακτηρίζουν όλο το διάστημα μελέτης. Τα πρότυπα αυτά μπορούν να χρησιμοποιηθούν, εκτός από τη «συμπίεση» της χρονοσειράς, για την κατασκευή αποτελεσματικών και στατιστικά εύρωστων (robust) μοντέλων βραχυπρόθεσμης πρόβλεψης.

Η Εικόνα 7 παρουσιάζει την ενεργειακή κατανομή των ICA παραγόντων (components), όπως υπολογίστηκαν για την ανάλυση του εβδομαδιαίου πίνακα αφίξεων (βλ. Εικόνα 1). Είναι φανερό ότι μόλις τρεις τέτοιοι παράγοντες, δηλαδή πρότυπα 7 ημερών, είναι αρκετά για την περιγραφή των 2/3 της διακύμανσης του σήματος. Επιπλέον, μέσω της ανάλυσης με αυτούς τους αλγορίθμους, επιβεβαιώθηκε και πάλι ότι για αποτελεσματική βραχυπρόθεσμη πρόβλεψη απαιτείται συνεχόμενο διάστημα προηγούμενης καταγραφής περίπου τριών εβδομάδων (21 ημερών).



Εικόνα 6: Φασματική απεικόνιση (DFT/FFT) των ημερήσιων αφίξεων.



Εικόνα 7: Ανάλυση παραγόντων (ICA) των εβδομαδιαίων αφίξεων.

### Συζήτηση

Τα αποτελέσματα της μελέτης αποδεικνύουν ότι η ανάλυση των ημερήσιων αφίξεων μπορεί πράγματι να χρησιμοποιηθεί για τη βραχυπρόθεσμη πρόβλεψη των προσφυγικών ροών.

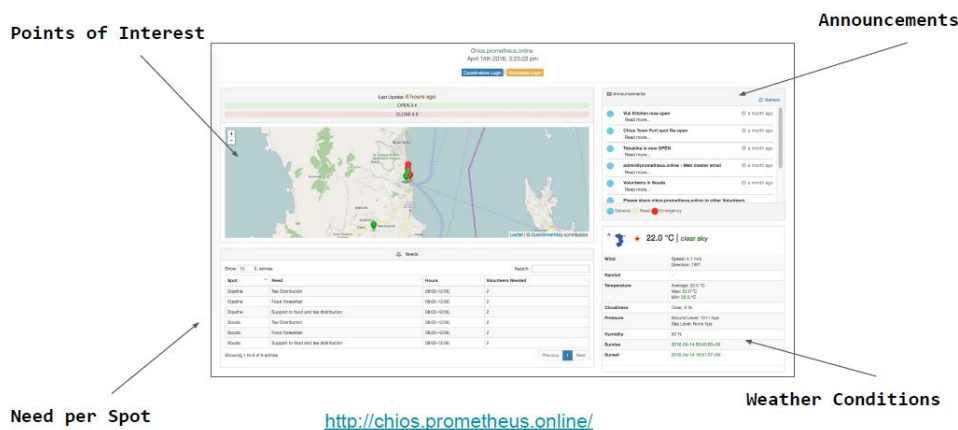
Μοντέλα αυτοσυσχέτισης-εξομάλυνσης (ARMA), καθώς και φασματικής ανάλυσης, μπορούν να εφαρμοστούν για την εκτίμηση του «επιπέδου κινδύνου» σε επιχειρήσεις έρευνας και διάσωσης, αντίστοιχα με τα συστήματα «επιπέδου επιφυλακής» που εφαρμόζονται π.χ. στην έγκαιρη προετοιμασία για την αντιμετώπιση πυρκαγιών κατά τους θερινούς μήνες.

Επιπλέον, η ανάλυση παραγόντων σε εβδομαδιαία βάση μπορεί να προσφέρει εξαιρετικά χρήσιμες πληροφορίες για τη συμπεριφορά των παράνομων δικτύων διακίνησης, καθώς και των ημερών έντονης δραστηριότητας, όχι απλά σε επίπεδο

στατιστικών, αλλά σε επίπεδο ενδογενών χαρακτηριστικών του συστήματος. Δεν είναι τυχαίο που η συμπεριφορά των δικτύων αυτών προσομοιάζει τη συμπεριφορά των δικτύων δεδομένων, όπου «πακέτα» διακινούνται μεταξύ κόμβων, αποθηκεύονται σε προσωρινούς χώρους «buffers» και τελικά μεταδίδονται ως εκροές.

Πιλοτική αξιοποίηση των παραπάνω μοντέλων πρόβλεψης αναπτύσσεται ήδη σε συνδυασμό με αντίστοιχες εφαρμογές λογισμικού για τον επιχειρησιακό συντονισμό (Emergency Operations Center - EOC) των εθελοντικών ομάδων σε Λέσβο («Sahana4Greece»), Χίο και Σάμο («Prometheus»).

Αν και μετά τη συμφωνία μεταξύ ΕΕ-Τουρκίας οι προσφυγικές ροές στα ελληνικά νησιά έχουν μειωθεί σε μεγάλο βαθμό, εντούτοις το πρόβλημα εν γένει παραμένει και είναι πιθανό να εμφανιστεί ξανά στο μέλλον. Ήδη, λόγω της μετατόπισης των προσφυγικών ροών και πάλι προς τα περάσματα στην κεντρική Μεσόγειο, το ίδιο ακριβώς πρόβλημα αντιμετωπίζει τώρα η Ιταλία στα νότια θαλάσσια σύνορά της, με πολύ υψηλότερο αναλογικά ποσοστό θανάτων (ως και 1 προς 30) λόγω της απόστασης και των πιο επικίνδυνων συνθηκών. Επομένως η ανάγκη για παρόμοια συστήματα έγκαιρης προειδοποίησης παραμένει πιο επίκαιρη από ποτέ.



Εικόνα 8: Η πλατφόρμα “Prometheus” για το συντονισμό των εθελοντικών ομάδων, όπως βρίσκεται σε επιχειρησιακή χρήση στη Χίο από τον Ιανουάριο 2016.

## Βιβλιογραφία

Georgiou H (2016) Identification of refugee influx patterns in Greece via model-theoretic analysis of daily arrivals. ArXiv:1605.02784 [stat.ML]

Anastasiadis S, Georgiou H (2016) Prometheus: Virtual Emergency Response Center – the Chios island deployment. FOSSCOMM 2016 @ 16-17 Apr 2016, Athens

Georgiou H (2015) Sahana4Greece: A crowd-sourced virtual EOC for supporting the rescue & relief operations in Greece for the refugees. FOSSCOMM 2015 @ 7-8 Nov 2015, Athens

S. Anastasiadis (2016) Prometheus: The virtual Emergency Operations Center for Chios [Online] <http://chios.prometheus.online>

UNHCR (2016) Mediterranean sea arrivals – 2015 data by location, country of arrival (data) [Online] <http://data.unhcr.org/mediterranean/download.php?id=490>