

Συγκριτική ανάλυση της οδηγικής επίδοσης με χρήση τεχνικών της επιστήμης των δεδομένων εφαρμοζόμενων σε δεδομένα κινητών τηλεφώνων μεγάλης κλίμακας



Δημήτρης Τσελέντης

Πολιτικός – Συγκοινωνιολόγος Μηχανικός
Υποψήφιος Διδάκτωρ – Ερευνητής

www.nrso.ntua.gr/dtsel/
dtsel@central.ntua.gr

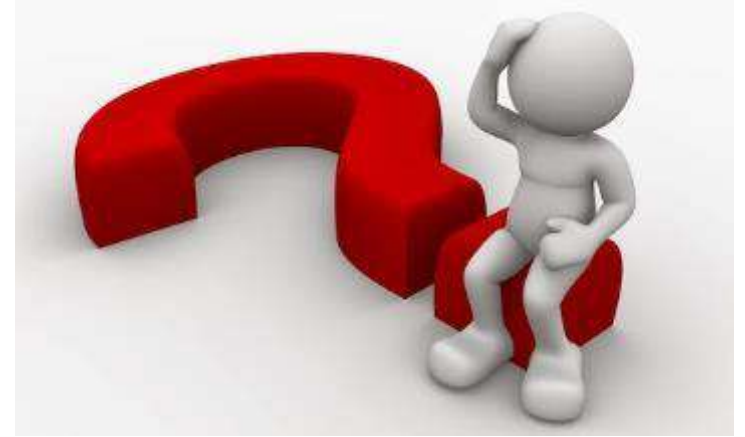
Σκοπός

- **Μεθοδολογική προσέγγιση** για την συγκριτική αξιολόγηση της οδηγικής επίδοσης από άποψη ασφαλείας:
 - διαδρομή
 - οδηγός
 - ανάλυση πολλαπλών κριτηρίων
- **Δείκτης** επίδοσης **ασφαλείας**
 - χαρακτηριστικά διαδρομής
 - χαρακτηριστικά οδηγικής συμπεριφοράς
 - κινητά τηλέφωνα
- «Έξυπνα» **κινητά** τηλέφωνα
 - δεδομένα ευρείας κλίμακας
 - πείραμα φυσικής οδήγησης



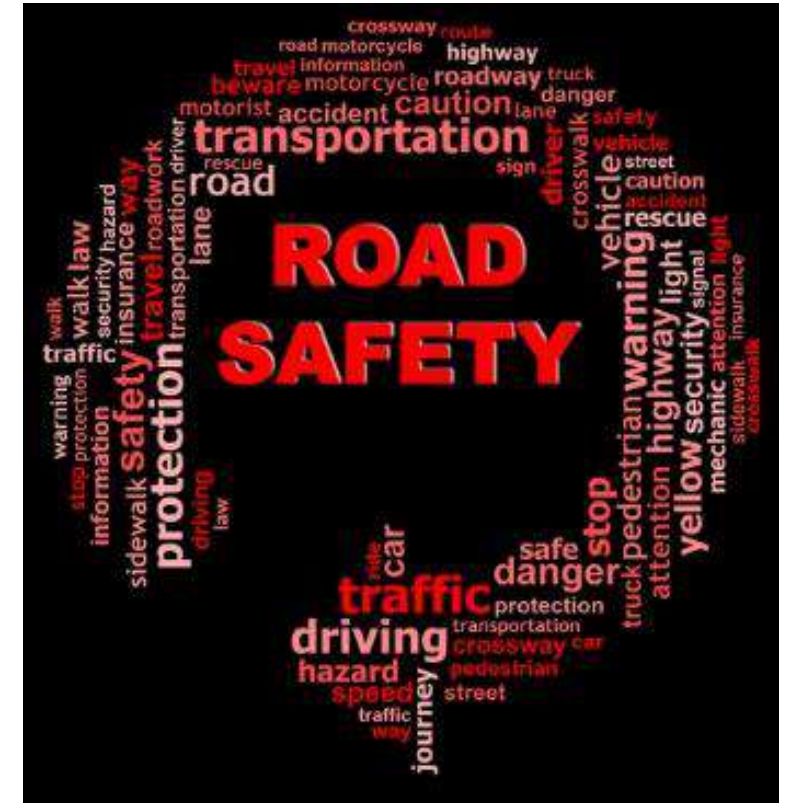
Ερευνητικά ερωτήματα

- Πόσο επαρκώς μπορεί να **αξιολογηθεί** η οδηγική **επίδοση** ως προς την ασφάλεια; Μπορούν οι τεχνικές της επιστήμης των δεδομένων να δώσουν ικανοποιητικές απαντήσεις;
- Ποια είναι τα χαρακτηριστικά της **εξέλιξης** της οδηγικής επίδοσης μέσα στο **χρόνο**; Τι αντιπροσωπεύουν οι ομάδες οδηγών που σχηματίζονται;
- Ποιο είναι το απαραίτητο **μέγεθος δεδομένων** που απαιτείται να συλλεχθεί για κάθε οδηγό;
- Πως μπορούν να εντοπιστούν οι **λιγότερο αποδοτικές** διαδρομές σε μια βάση δεδομένων;



Ανασκόπηση βιβλιογραφίας (1/4)

- Μεθοδολογίες μέτρησης της **ασφαλούς** οδηγικής συμπεριφοράς
 - μικρός αριθμός παραμέτρων
 - συσχέτιση μεταξύ των παραγόντων
 - ποσοτικοποίηση της συνολικής οδηγικής επίδοσης
 - πολυκριτηριακή ανάλυση
- **Γραμμικός προγραμματισμός** για μέτρηση επίδοσης
 - Περιβάλλουσα Ανάλυση Δεδομένων (DEA)
 - Τράπεζες, Επιχειρήσεις, Νοσοκομεία, Προσωπικό, κτλ.
 - Μεταφορές (Συστήματα μεταφορών, Οδική ασφάλεια κτλ..)
 - Έρευνα οδηγικής συμπεριφοράς



Ανασκόπηση βιβλιογραφίας (2/4)

- **Εξάλειψη των εμποδίων** που υπήρχαν:
 - Τεχνολογία κινητών τηλεφώνων
 - Υψηλό κόστος
 - συστημάτων καταγραφής εντός του οχήματος (π.χ. OBD)
 - προγραμμάτων δεδομένων
 - cloud computing
 - Χαμηλός ρυθμός διείσδυσης των «έξυπνων» κινητών τηλεφώνων
 - Ανεπαρκής διαχείριση και αξιοποίηση των δεδομένων ευρείας κλίμακας
- Παρούσα **τεχνολογική εξέλιξη**
 - συλλογή και αξιοποίηση των δεδομένων μέσω κινητών τηλεφώνων
 - ευκολότερα και με μεγαλύτερη ακρίβεια



Ανασκόπηση βιβλιογραφίας (3/4)

- Συλλογή **δεδομένων οδήγησης**
 - Πειράματα φυσικής οδήγησης
 - Πειράματα σε προσομοιωτή οδήγησης
 - Σε βάθος ανάλυση ατυχημάτων
- Χαρακτηριστικά οδήγησης – Επαρκής **όγκος**
 - αξιολόγηση κάθε χρήστη
 - ανεπαρκής όγκος δεδομένων => αβέβαια ή αδικαιολόγητα αποτελέσματα
 - υπερβολικός όγκος δεδομένων => σημαντική αύξηση του απαιτούμενου υπολογιστικού χρόνου



Ανασκόπηση βιβλιογραφίας (4/4)

- Ασφάλιση βασισμένη στη χρήση (**UBI**)
 - Pay-As-You-Drive (PAYD)
 - Pay-How-You-Drive (PHYD)
 - Pay-at-the-pump (PATP)
- Χαρακτηριστικά **διαδρομής**
 - Συνολική απόσταση
 - Τύπος οδικού δικτύου
 - Επικίνδυνες ώρες οδήγησης
 - Συχνότητα διαδρομής
 - Τύπος οχήματος
 - Καιρικές συνθήκες
- Χαρακτηριστικά **οδηγικής συμπεριφοράς**
 - Υπέρβαση ορίου ταχύτητας
 - Απότομες επιβραδύνσεις/ επιταχύνσεις/ στροφές
 - Χρήση ζώνης ασφαλείας
 - Χρήση κινητού τηλεφώνου



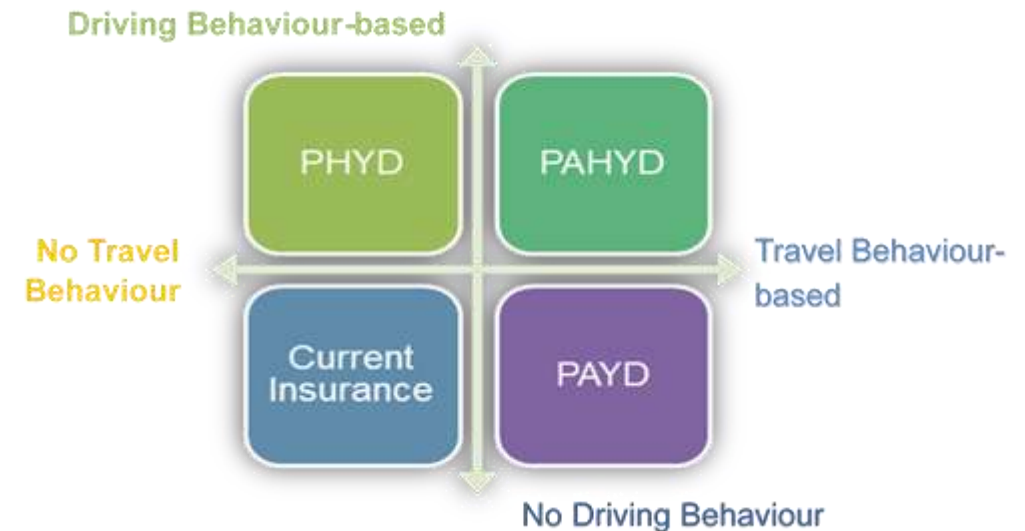
Βιβλιογραφικό κενό (1/2)

- **Συγκριτική ανάλυση** της οδηγικής επίδοσης ως προς την ασφάλεια
 - μικροσκοπικά δεδομένα οδήγησης
 - χαρακτηριστικά διαδρομής και οδήγησης
- Δεδομένα **ευρείας κλίμακας** από πειράματα φυσικής οδήγησης
- Χαρακτηριστικά **ανθρώπινου παράγοντα** καταγεγραμμένα από κινητά τηλέφωνα
 - Απότομες επιταχύνσεις/ επιβραδύνσεις
 - Χρόνος χρήσης κινητού τηλεφώνου
 - Χρόνος υπέρβασης του ορίου ταχύτητας



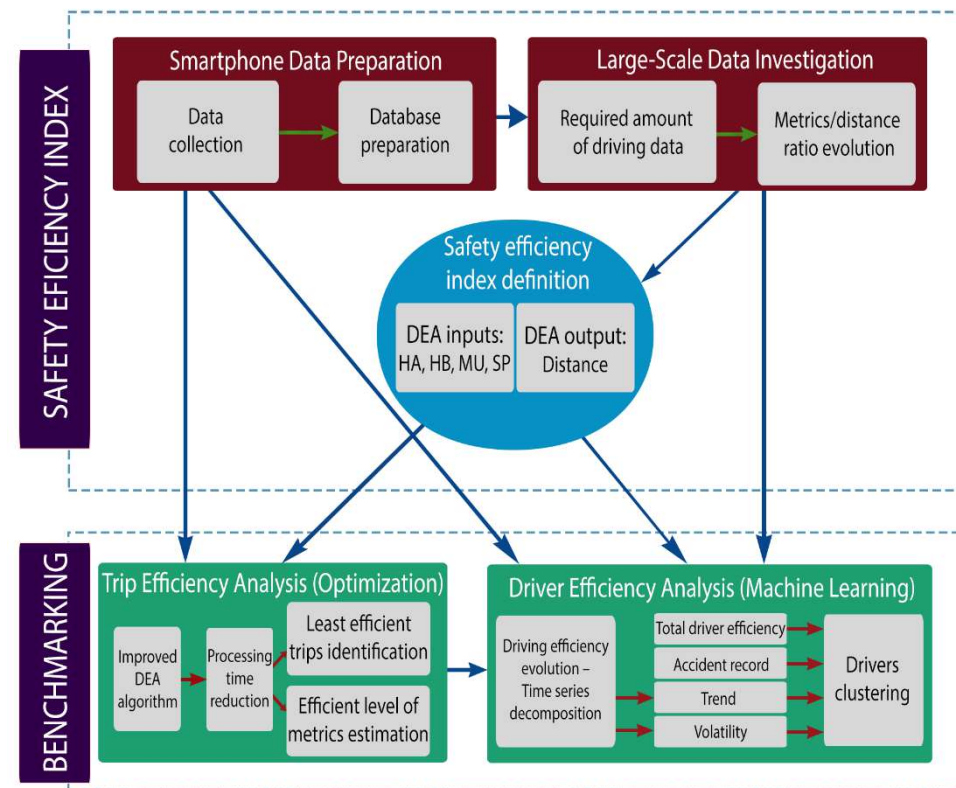
Βιβλιογραφικό κενό (2/2)

- **Ποσότητα** δεδομένων οδήγησης
- Σχήματα ασφάλισης βασισμένα στη **χρήση**
 - χαρακτηριστικά διαδρομής και οδήγησης
 - ετήσια διανυόμενη χιλιομετρική απόσταση
 - οδηγική συμπεριφορά
- **Παράγοντες επικινδυνότητας**
 - πιθανότητα ατυχήματος
 - οδηγική συμπεριφορά
 - ετήσια διανυόμενη χιλιομετρική απόσταση



Μεθοδολογική προσέγγιση

- **Συλλογή** δεδομένων από κινητά τηλέφωνα
 - προετοιμασία της βάσης δεδομένων
- **Διερεύνηση** δεδομένων μεγάλης κλίμακας
 - μελέτη της εξέλιξης των χαρακτηριστικών οδήγησης
 - επαρκής ποσότητα δεδομένων οδήγησης
- Προσδιορισμός **δείκτη** οδηγικής **ασφάλειας**
- Ανάλυση **επίδοσης** σε επίπεδο **διαδρομής**
 - εντοπισμός των λιγότερο αποδοτικών διαδρομών
 - εκτίμηση του αποδοτικού επιπέδου των χαρακτηριστικών οδήγησης
- Ανάλυση **επίδοσης** σε επίπεδο **οδηγού**
 - μελέτη της χρονικής εξέλιξης
 - συσταδοποίηση οδηγών



Περιβάλλουσα ανάλυση δεδομένων (1/2)

- Τεχνική **βελτιστοποίησης**
 - Μέτρηση της απόδοσης με χρήση **DEA**
 - Εταιρείες, τράπεζες, νοσοκομεία, προσωπικό κτλ.
- **Σημαντικά υψηλό** υπολογιστικό κόστος
 - ακριβής λύση
 - τεχνική μειούμενων δεδομένων εισόδου (RBE DEA)
 - Convex Hull DEA
 - μείωση χρόνου υπολογισμού
- **Μονάδα απόφασης (DMU)**
 - Εργοστάσιο, εταιρεία κτλ.
 - Διαδρομές, οδηγοί
 - συνεχείς και ποσοτικές μεταβλητές
 - κάθε οδηγός θα πρέπει να μειώσει την συχνότητα των οδηγικών χαρακτηριστικών του για δεδομένη απόσταση



Περιβάλλουσα ανάλυση δεδομένων (2/2)

- **Δείκτης επίδοσης** $\text{Driving_Efficiency}_B$
- **Input-oriented** DEA
 - ελαχιστοποίηση των inputs (αριθμός συμβάντων HA, HB κτλ.) ανά μονάδα απόστασης
- Πρόβλημα **constant-returns-to-scale** (CRS)
 - το άθροισμα των inputs αλλάζει αναλογικά με το άθροισμα των outputs (απόσταση)
- **Αποδοτικό επίπεδο** διαδρομής/ οδηγού
 - inputs
 - outputs

$$\min(\text{Driving_Efficiency}_B)$$

Subject to the following constraints:

$$\text{Driving_Efficiency}_B * x_o - X * \lambda \geq 0$$

$$Y * \lambda \geq y_o$$

$$\lambda_i \geq 0 \forall \lambda_i \in \lambda$$

$$\text{Metric}_i = \sum_{j=1}^m \lambda_j * \text{Metric}_j$$

$$\text{distance}_{\text{urban}} = \text{distance}_i / \text{Driving_Efficiency}_i$$

Παράμετροι δείκτη επίδοσης

Παράμετροι έκθεσης κινδύνου:

- Συνολική διανυθείσα **απόσταση**

Παράμετροι οδηγικής συμπεριφοράς:

- **Απότομα** συμβάντα
 - αριθμός απότομων επιβραδύνσεων (διαμήκης επιβράδυνση) (HB)
 - αριθμός απότομων επιταχύνσεων (διαμήκης επιτάχυνση) (HA)
- Υπέρβαση **ορίου ταχύτητας** (SP)
 - δευτερόλεπτα υπέρβασης ορίου ταχύτητας
- Απόσπαση προσοχής από χρήση **κινητού τηλεφώνου** (MU)
 - δευτερόλεπτα χρήσης κινητού τηλεφώνου

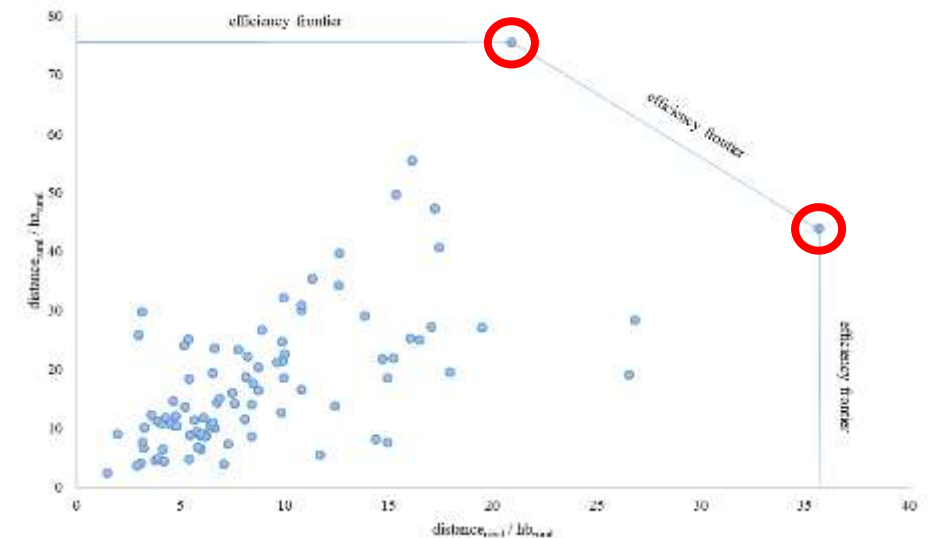
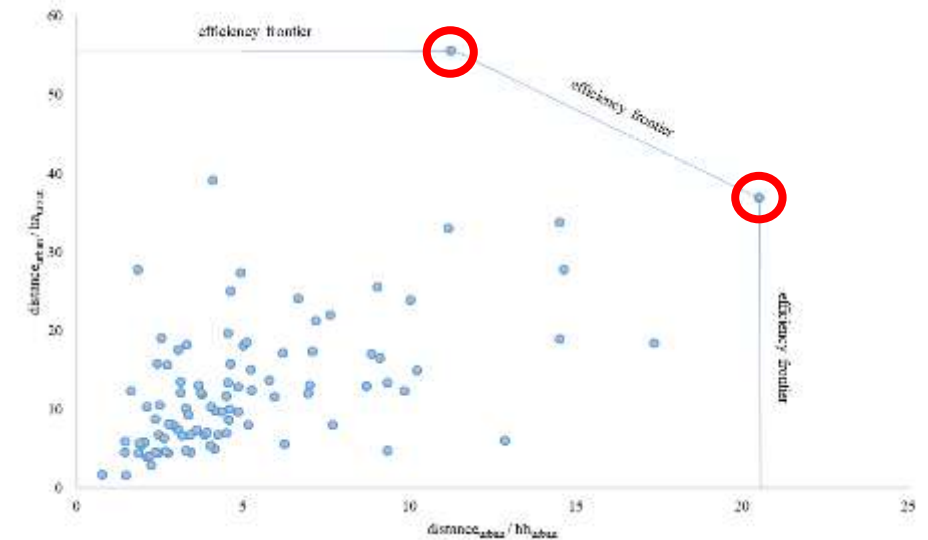
Τύποι οδών

- Αστικό οδικό δίκτυο (σηματοδοτούμενο ή μη, ταχύτητα ≤ 50 χλμ/ώρα)
- Αστικό οδικό δίκτυο ταχείας κυκλοφορίας (ταχύτητα 50 - 90 χλμ/ώρα)
- Υπεραστικό δίκτυο (ταχύτητα ≥ 90 χλμ/ώρα)



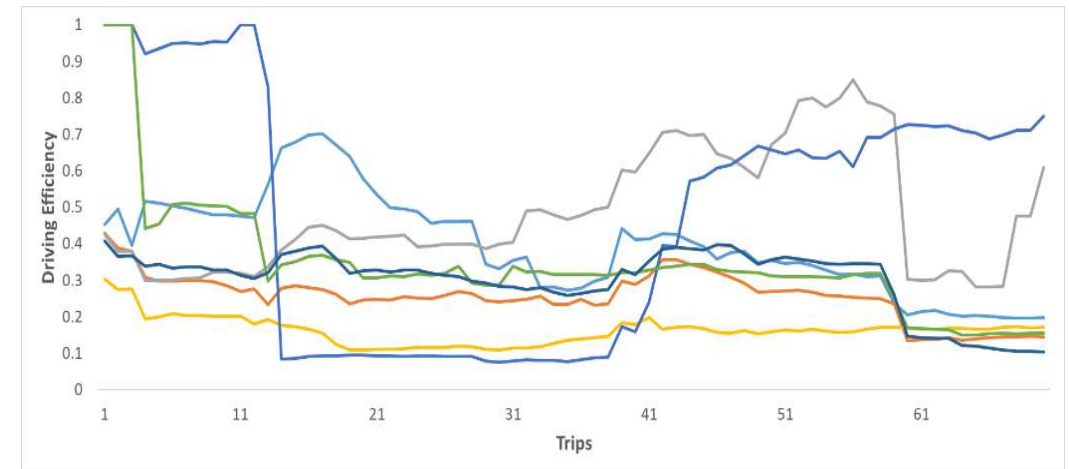
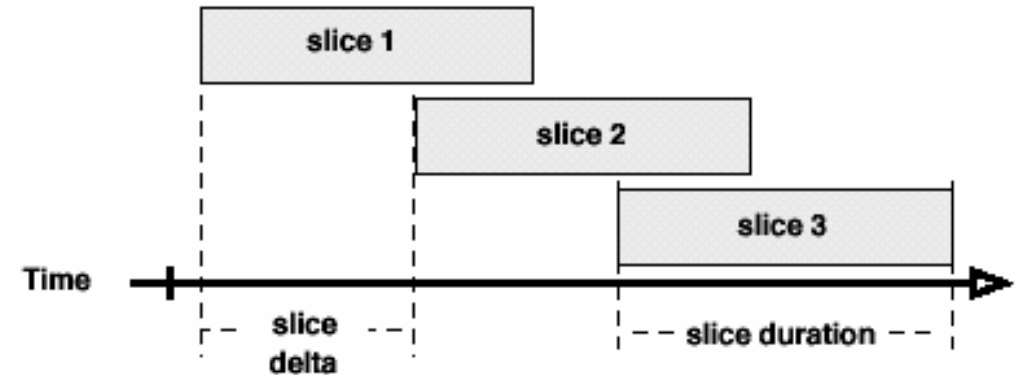
Ανάλυση οδηγικής επίδοσης με χρήση DEA

- **Απεικόνιση 2-D** των αποτελεσμάτων της DEA
 - Αστικό οδικό δίκτυο
 - Αστικό οδικό δίκτυο ταχείας κυκλοφορίας
- **DEA inputs**
 - αριθμός συμβάντων HA
 - αριθμός συμβάντων HB
- **DEA outputs**
 - απόσταση διαδρομής



Χρονική εξέλιξη επίδοσης

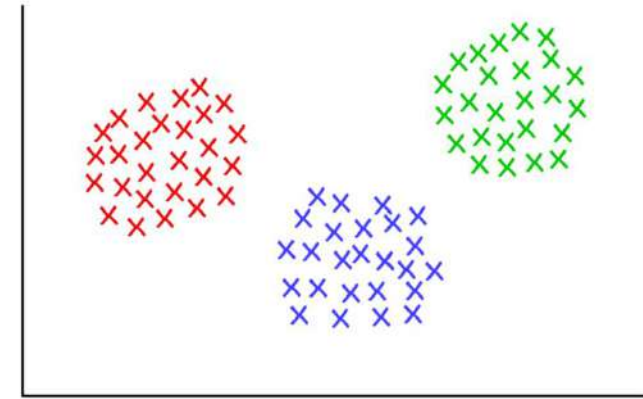
- Συγκριτική **ανάλυση** της οδηγικής επίδοσης **συγκεντρωτικά**
- Συγκριτική **ανάλυση** της οδηγικής επίδοσης σε **κυλιόμενο παράθυρο**
 - μεταβολές οδηγικής συμπεριφοράς
 - ανάλυση χρονοσειρών
 - αποσύνθεση των χρονοσειρών
 - Στασιμότητα
 - Τάση
 - Μεταβλητότητα
 - τύπος οδού



Συσταδοποίηση οδηγών

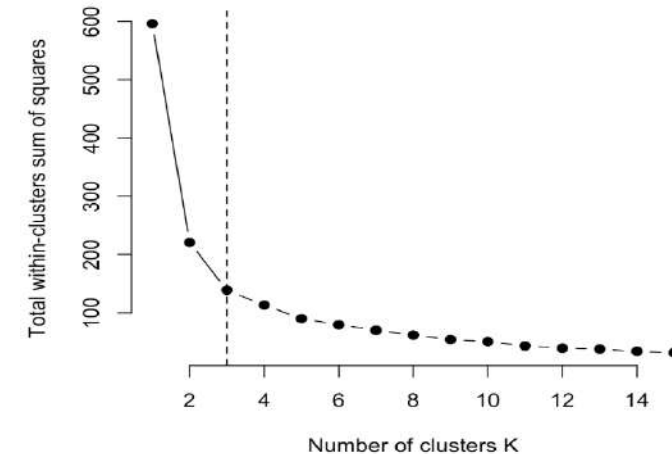
- Συσταδοποίηση οδηγών

- οδηγικά προφίλ και τα χαρακτηριστικά τους
- αλγόριθμος k-means
- μέθοδος του «αγκώνα»



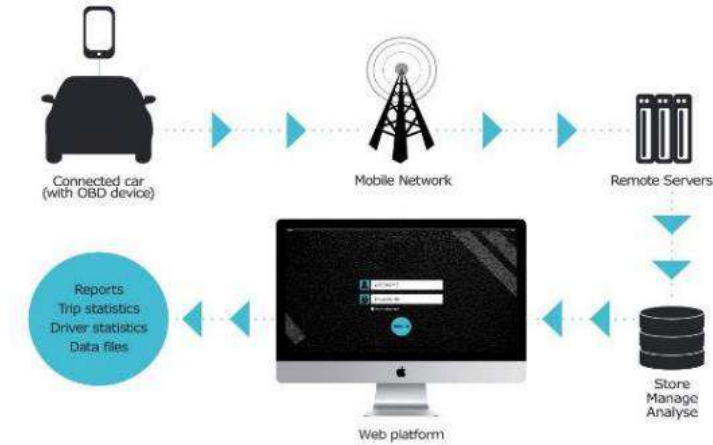
- Χαρακτηριστικά συσταδοποίησης

- συνολική οδηγική επίδοση
- τάση
- μεταβλητότητα
- χωρίς ιστορικό ατυχημάτων (data_sample_1)
- με ιστορικό ατυχημάτων (data_sample_2)
- σύγκριση αποτελεσμάτων



Δεδομένα κινητών τηλεφώνων

- **Εφαρμογή** κινητού τηλεφώνου για την αυτόματη καταγραφή της συμπεριφοράς του χρήστη
- Χρήση **APIs** για την ανάγνωση των δεδομένων από τους αισθητήρες των κινητών τηλεφώνων
- **Μετάδοση** δεδομένων από την εφαρμογή στην κεντρική βάση δεδομένων
- **Αποθήκευση** των δεδομένων στη βάση δεδομένων για τη διαχείριση και επεξεργασία τους
- Προσδιορισμός οδηγικών **χαρακτηριστικών** με χρήση
 - αλγορίθμων μηχανικής μάθησης
 - τεχνικών εξόρυξης δεδομένων ευρείας κλίμακας



Πηγή: OSeven Telematics

Δεδομένα ερωτηματολογίου

- **Οδηγική εμπειρία**
 - έτη οδήγησης
 - ετήσια διανυόμενη απόσταση
- **Όχημα**
 - ηλικία οχήματος
 - κυβισμός
 - ιδιοκτησία
 - κατανάλωση καυσίμων
- **Οδηγική συμπεριφορά**
 - ιστορικό ατυχημάτων
 - αυτο-αξιολόγηση
- **Δημογραφικά χαρακτηριστικά**
 - ηλικία
 - φύλο
 - εκπαίδευση



Προετοιμασία δεδομένων

- **Ανωνυμοποίηση** δεδομένων

- προσέγγιση αγνωστικιστική ως προς τον χρήστη
- αναγνώριση οδηγικών συμπεριφορών και μοτίβων
- αιτιότητα μεταξύ της συμπεριφοράς και άλλων χαρακτηριστικών
- δείγματα μεγάλης κλίμακας
- χωρίς προηγούμενη πληροφόρηση για τα δημογραφικά χαρακτηριστικά ή του ιστορικού ατυχημάτων



- Γλώσσα προγραμματισμού **Python**

- φιλτράρισμα
- διατήρηση μόνο των απαραίτητων πληροφοριών
- συγκεντρωτικά δεδομένα
- ανάλυση δεδομένων



Δείγμα οδηγών

	Sampling time investigation	Trip efficiency analysis	Driver efficiency analysis			
			data_sample_1		data_sample_2	
			Urban	Rural	Urban	Rural
Number of drivers	171	88	100	100	43	39
Number of trips	49,722	10,088	23,000	15,000	9,890	5,850

Δείγμα οδηγών - Ανάλυση οδηγών

- Κριτήρια

- τουλάχιστον 50 περισσότερες διαδρομές
 - ελάχιστος απαιτούμενος αριθμός
 - δημιουργία των χρονοσειρών
- μη μηδενική απόσταση σε όλους τους τύπους οδών
- μη μηδενικό άθροισμα των δεδομένων εισόδου της DEA (HA, HB κτλ.)
- μέγιστος αριθμός οδηγών



- Αστικό οδικό δίκτυο

- 100 οδηγοί & 23,000 διαδρομές – χωρίς δεδομένα ερωτηματολογίου
- 43 οδηγοί & 9,890 διαδρομές – με δεδομένα ερωτηματολογίου

- Αστικό οδικό δίκτυο ταχείας κυκλοφορίας

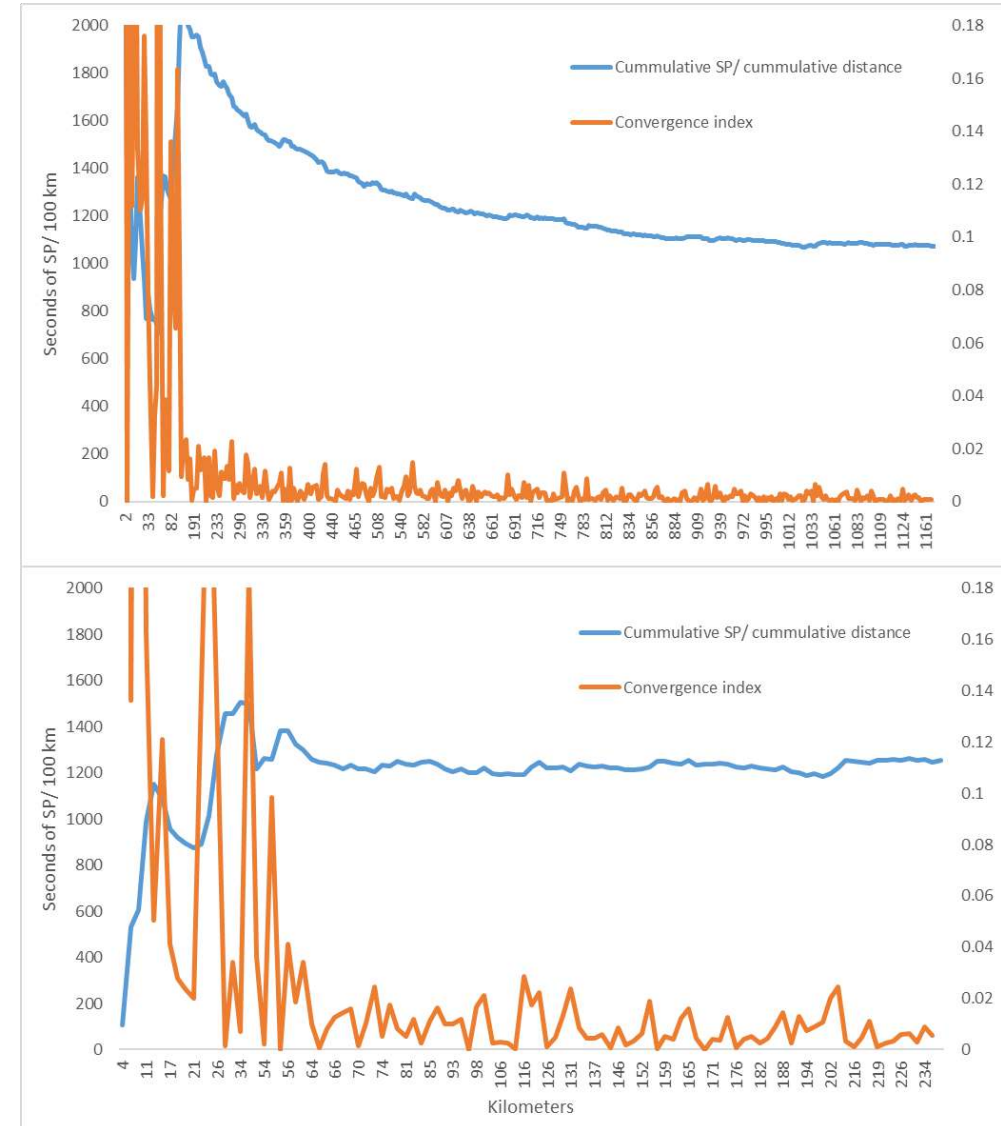
- 100 οδηγοί & 15,000 διαδρομές – χωρίς δεδομένα ερωτηματολογίου
- 39 οδηγοί & 5,850 διαδρομές – με δεδομένα ερωτηματολογίου



Διερεύνηση δεδομένων (1/7)

- Δείκτης σύγκλισης

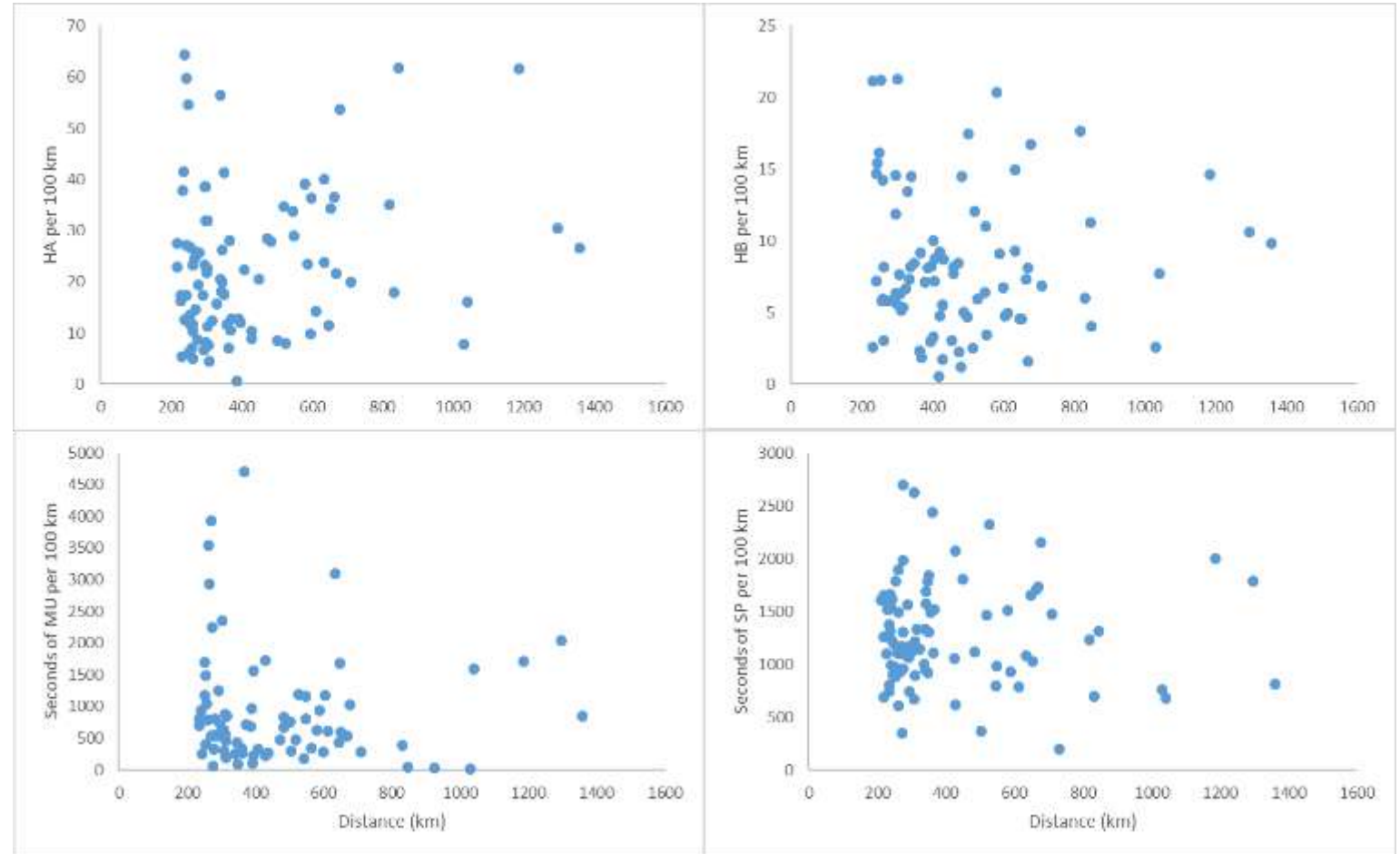
- κυλιόμενο «παράθυρο» 40 διαδρομών
- τουλάχιστον 200 χλμ
- $\leq 5\%$



Διερεύνηση δεδομένων (2/7)

Αστικό οδικό δίκτυο

- Ασθενής **θετική συσχέτιση** μεταξύ των **HA** και της απαιτούμενης απόστασης καταγραφής
- Η απαιτούμενη απόσταση καταγραφής είναι **υψηλότερη** για τους **πιο** επιθετικούς οδηγούς
- **Καμία** προφανής **τάση** για τα υπόλοιπα χαρακτηριστικά



Διερεύνηση δεδομένων (3/7)

Αστικό οδικό δίκτυο

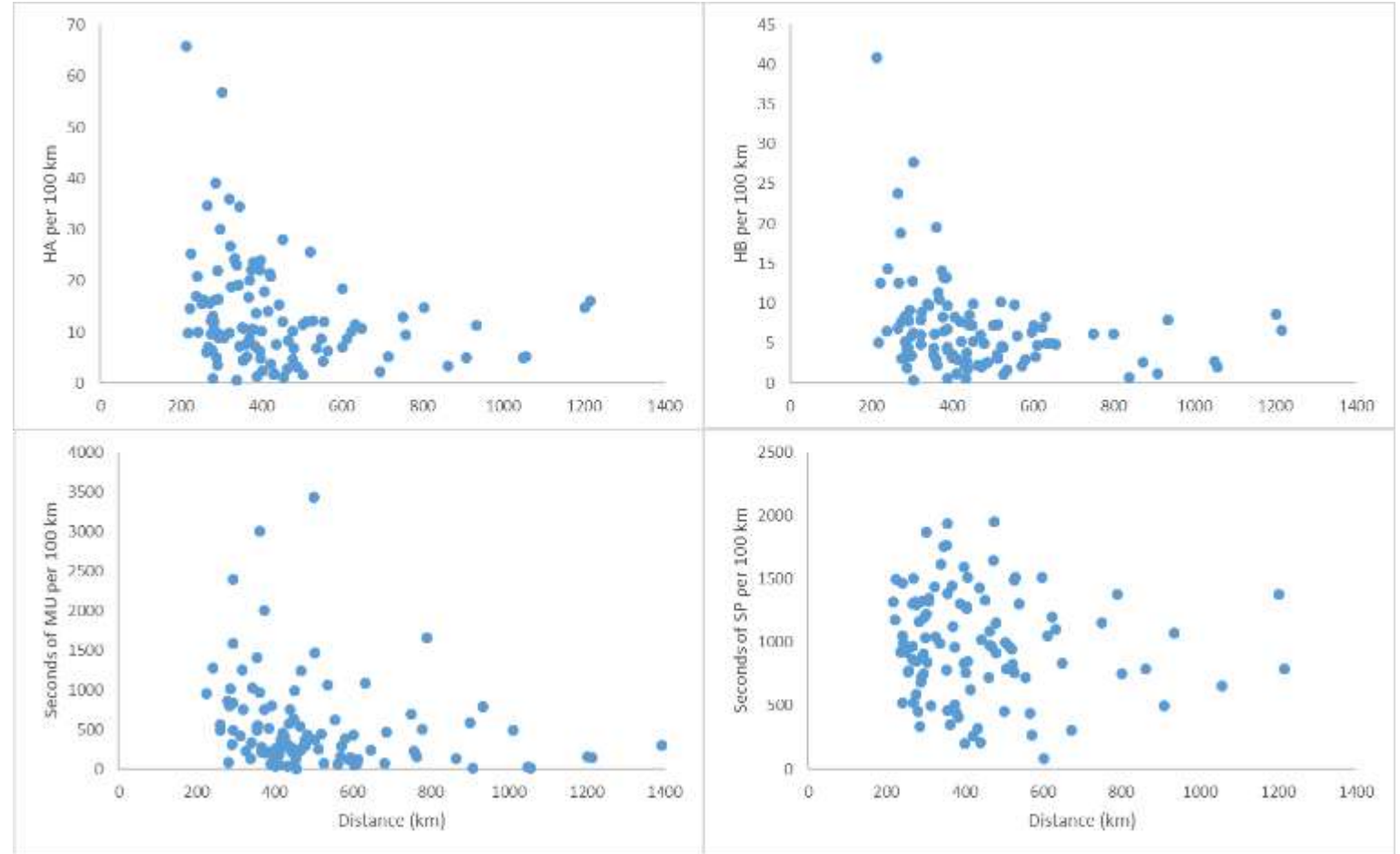
- **Ίδια** απαιτούμενη απόσταση **δειγματοληψίας** για όλα τα επίπεδα επιθετικότητας/ επικινδυνότητας
- **Μέγιστη** διάμεσος της **απόστασης**
 - Σύγκλιση όλων των χαρακτηριστικών στον αθροιστικό μ.ο.
 - το οδηγικό δείγμα είναι επαρκές
 - ο λόγος των δεδομένων εισόδου προς τα δεδομένα εξόδου είναι σταθερός (απαίτηση DEA)
- **519** χλμ.
 - 75 διαδρομές

Metric	Percentile range	Metric descriptive statistics				Distance to convergence		
		Average	St. Dev	Min	Max	Average	Median	St. Dev
HA	0% – 25%	8.18	2.61	-	11.61	389	309	177
	25% – 50%	15.92	2.8	11.61	20.54	399	336	208
	50% – 75%	25.01	2.48	20.54	30.14	415	322	246
	75% – 100%	43.23	10.92	30.14	-	526	519	293
HB	0% – 25%	3.05	1.26	-	4.95	515	478	179
	25% – 50%	6.17	0.69	4.95	7.32	408	335	155
	50% – 75%	8.6	0.83	7.32	10.61	558	431	300
	75% – 100%	15.65	3.12	10.61	-	469	341	248
MU	0% – 25%	204	101	-	332	559	407	395
	25% – 50%	495	78	332	606	440	347	174
	50% – 75%	799	111	606	1041	443	381	247
	75% – 100%	2063	994	1041	-	493	366	310
SP	0% – 25%	727	194	-	947	489	339	308
	25% – 50%	1081	67	947	1198	343	293	123
	50% – 75%	1402	119	1198	1594	378	312	188
	75% – 100%	1919	318	1594	-	455	348	287

Διερεύνηση δεδομένων (4/7)

Αστικό οδικό δίκτυο ταχείας κυκλοφορίας

- Ασθενής **αρνητική** συσχέτιση μεταξύ των **HA, HB**, χρήσης **κινητού τηλεφώνου** και της απαιτούμενης απόστασης καταγραφής
- Η απαιτούμενη απόσταση καταγραφής είναι **υψηλότερη** για τους **λιγότερο** επιθετικούς/ επικίνδυνους/ αφηρημένους οδηγούς
- **Καμία** προφανής **τάση** για την υπέρβαση ορίου ταχύτητας



Διερεύνηση δεδομένων (5/7)

Αστικό οδικό δίκτυο ταχείας κυκλοφορίας

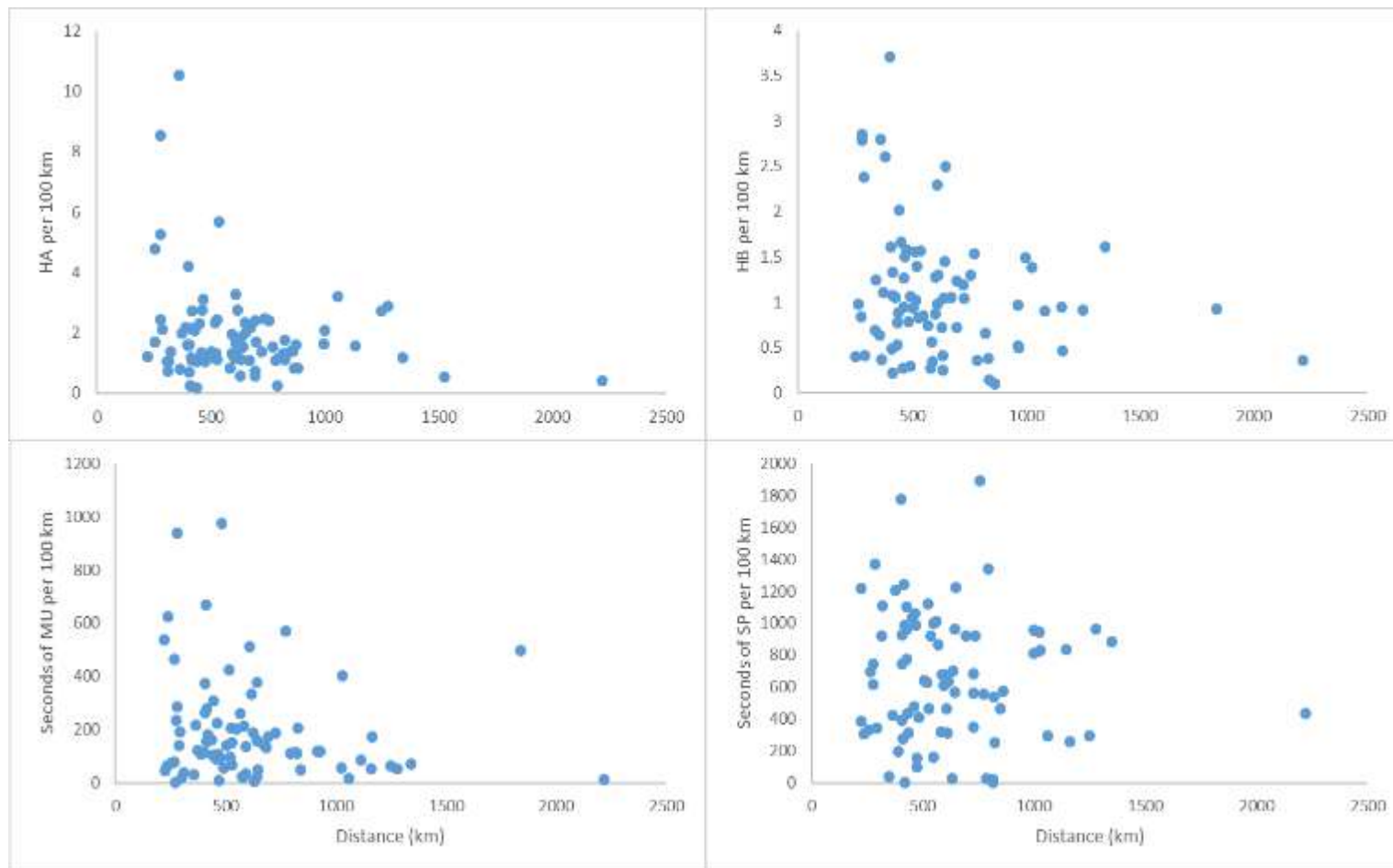
- **Διαφορετική** απαιτούμενη απόσταση **δειγματοληψίας** για κάθε επίπεδο επιθετικότητας/επικινδυνότητας
- **Μέγιστη** διάμεσος της απόστασης
- **579** χλμ.
 - 81 διαδρομές

Metric	Percentile range	Metric descriptive statistics				Distance to convergence		
		Average	St. Dev	Min	Max	Average	Median	St. Dev
HA	0% – 25%	3.69	1.7	-	6.42	507	431	219
	25% – 50%	8.8	1.33	6.42	10.79	405	376	134
	50% – 75%	13.58	1.98	10.79	17.02	499	417	270
	75% – 100%	27.27	11.04	17.02	-	349	336	85
HB	0% – 25%	2.05	0.84	-	3.15	520	439	221
	25% – 50%	4.35	0.67	3.15	5.6	429	403	126
	50% – 75%	6.92	0.69	5.6	8.28	489	440	221
	75% – 100%	13.54	7.16	8.28	-	389	360	187
MU	0% – 25%	85	48	-	157	630	579	261
	25% – 50%	263	50	157	371	485	442	209
	50% – 75%	511	92	371	747	501	448	192
	75% – 100%	1334	684	747	-	411	362	166
SP	0% – 25%	454	170	-	745	448	399	189
	25% – 50%	851	74	745	970	444	397	221
	50% – 75%	1142	112	970	1315	419	388	172
	75% – 100%	1526	181	1315	-	425	356	195

Διερεύνηση δεδομένων (6/7)

Υπεραστικό δίκτυο

- Ασθενής **αρνητική** συσχέτιση μεταξύ των **HA, HB**, χρήσης **κινητού τηλεφώνου** και της απαιτούμενης απόστασης καταγραφής
- Η απαιτούμενη απόσταση καταγραφής είναι **υψηλότερη** για τους **λιγότερο** επιθετικούς/ επικίνδυνους/ αφηρημένους οδηγούς
- **Καμία** προφανής **τάση** για την υπέρβαση ορίου ταχύτητας



Διερεύνηση δεδομένων (7/7)

Υπεραστικό δίκτυο

- **Διαφορετική** απαιτούμενη απόσταση **δειγματοληψίας** για κάθε επίπεδο επιθετικότητας/επικινδυνότητας
- **Μέγιστη** διάμεσος της απόστασης
- **611** χλμ
 - δεν διερευνάται

Metric	Percentile range	Metric descriptive statistics				Distance to convergence		
		Average	St. Dev	Min	Max	Average	Median	St. Dev
HA	0% – 25%	0.74	0.29	-	1.1	678	585	440
	25% – 50%	1.26	0.12	1.1	1.54	639	605	234
	50% – 75%	1.87	0.24	1.54	2.3	607	611	242
	75% – 100%	3.77	2.12	2.3	-	593	529	289
HB	0% – 25%	0.36	0.12	-	0.56	705	592	413
	25% – 50%	0.83	0.1	0.56	0.97	700	562	376
	50% – 75%	1.15	0.13	0.97	1.38	572	606	174
	75% – 100%	2.05	0.64	1.38	-	542	472	251
MU	0% – 25%	35	20	-	66	722	592	468
	25% – 50%	101	18	66	135	623	498	285
	50% – 75%	174	26	135	223	564	540	196
	75% – 100%	455	206	223	-	533	445	349
SP	0% – 25%	193	124	-	346	610	550	285
	25% – 50%	505	91	346	641	649	592	392
	50% – 75%	807	100	641	950	666	601	295
	75% – 100%	1168	249	950	-	544	464	241

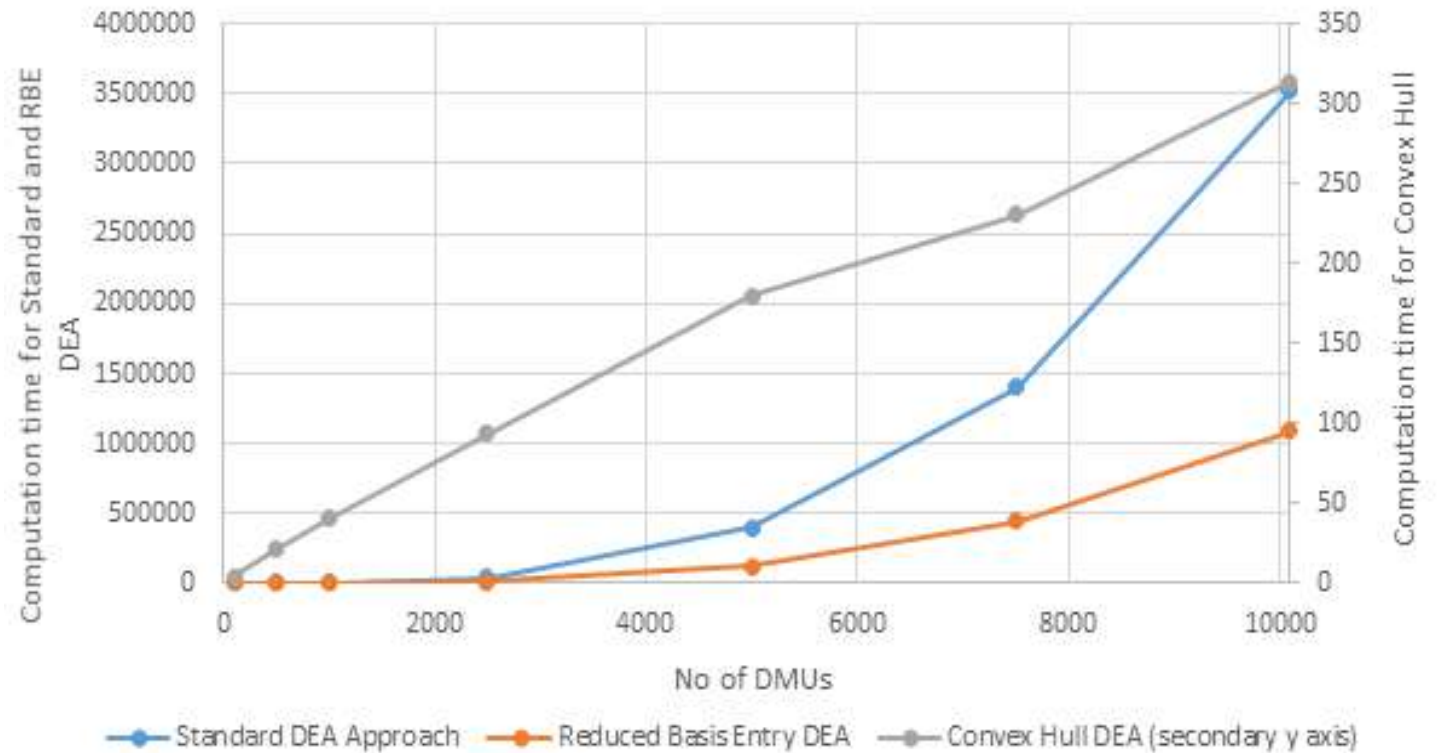
Ανάλυση διαδρομών (1/3)

- Η τεχνική του **Convex Hull** υπερέχει των άλλων 2
- Σημαντική **μείωση** του **χρόνου** υπολογισμού
 - Καθώς η βάση δεδομένων μεγαλώνει
 - Convex Hull DEA – 5 λεπτά
 - RBE DEA – 12.6 ημέρες
 - Standard DEA – 40.7 ημέρες

No of DMUs	Computation time (sec)			CH DEA % computation time improvement over		RBE DEA % computation time improvement over Standard DEA
	Standard DEA Approach	RBE DEA	Convex Hull DEA	Standard DEA Approach	RBE DEA	
100	11	6	4	63.64%	33.33%	45.45%
500	477	169	21	95.60%	87.57%	64.57%
1000	3250	1121	41	98.74%	96.34%	65.51%
2500	44435	15570	94	99.79%	99.40%	64.96%
5000	398485	123986	180	99.95%	99.85%	68.89%
7500	1400909	444498	231	99.98%	99.95%	68.27%
10088	3519372	1089731	314	99.99%	99.97%	69.04%
*Inputs = [ha _{urban} , ha _{rural} , ha _{highway}], Outputs = [distance _{urban} , distance _{rural} , distance _{highway}]						

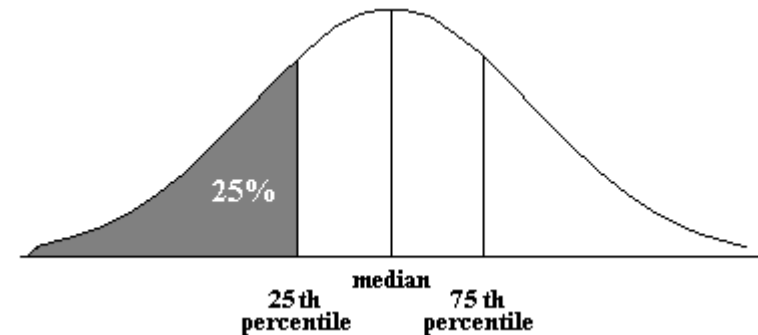
Ανάλυση διαδρομών (2/3)

- Convex Hull DEA
 - γραμμική αύξηση
- Πρότυπη DEA και RBE DEA
 - εκθετική αύξηση



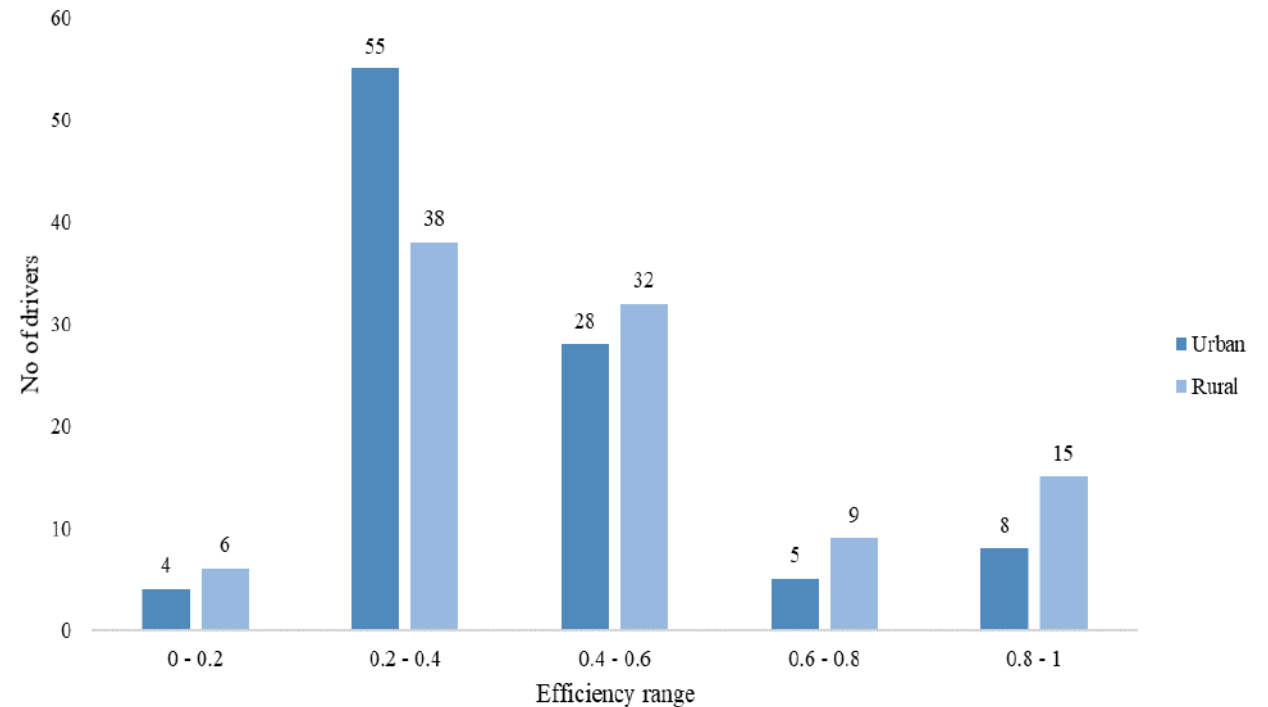
Ανάλυση διαδρομών (3/3)

- Μη αποδοτικές διαδρομές
- Εκτίμηση της επίδοσης
- Ταξινόμηση
 - μεγαλύτερο στο μικρότερο
- Επιλογή εκατοστημορίου
 - 5%, 10%, 25% κτλ.



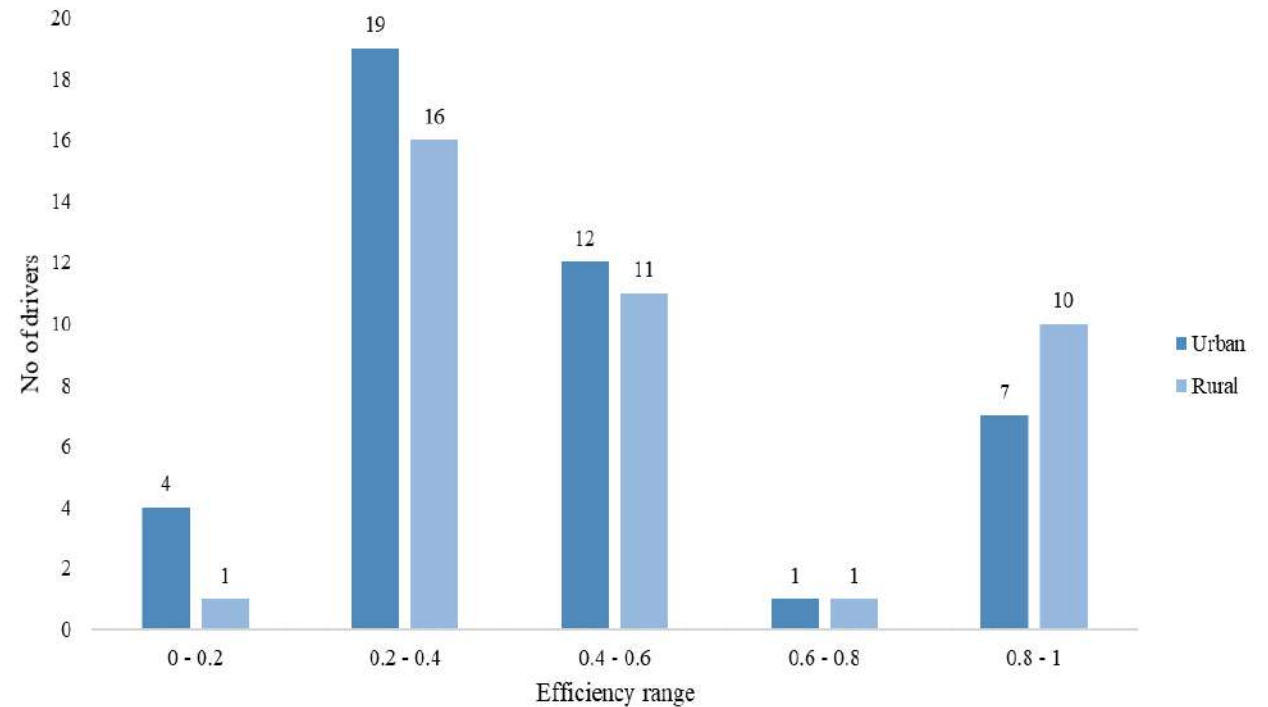
Ανάλυση οδηγών – Συγκεντρωτική επίδοση (1/13)

- data_sample_1
- **Παρόμοια κατανομή** και στους 2 τύπους οδών
- Συγκέντρωση οδηγών σε **χαμηλά-μεσαία** εύρη επίδοσης
 - η DEA αποδίδει έναν συγκριτικό δείκτη



Ανάλυση οδηγών – Συγκεντρωτική επίδοση (2/13)

- data_sample_2
- **Παρόμοια κατανομή** και στους 2 τύπους οδών
- Συγκέντρωση οδηγών σε **χαμηλά-μεσαία** εύρη επίδοσης
 - η DEA αποδίδει έναν συγκριτικό δείκτη



Ανάλυση οδηγών – Συγκεντρωτική επίδοση (3/13)

- **Συγκεντρωτική** οδηγική συμπεριφορά

- Εκατοστημότητα της οδηγικής επίδοσης
- 3 κατηγορίες επίδοσης
 - 1: Μη αποδοτικοί οδηγοί
 - 2: Ασθενώς αποδοτικοί οδηγοί
 - 3: Αποδοτικοί οδηγοί
- Διάμεσος οδηγικών χαρακτηριστικών

- **Αστικό οδικό δίκτυο**

- κατηγορία 1 -> κατηγορία 2
 - αριθμός απότομων επιταχύνσεων/ επιβραδύνσεων
- κατηγορία 2 -> κατηγορία 3
 - χρόνος υπέρβασης ορίου ταχύτητας

- **Αστικό οδικό δίκτυο ταχείας κυκλοφορίας**

- κατηγορία 1 -> κατηγορία 2
 - χρόνος υπέρβασης ορίου ταχύτητας
- κατηγορία 2 -> κατηγορία 3
 - αριθμός απότομων επιταχύνσεων
 - χρόνος χρήσης κινητού τηλεφώνου

Sample type	Road type	No of drivers	Driving characteristics (/100km)	Efficiency classes		
				Class 1: 0 - 25 % percentile	Class 2: 25 - 75 % percentile	Class 3: 75 - 100 % percentile
data_sample_1	Urban	100	efficiency	0.22	0.36	0.61
			ha	21.49	11.82	8.82
			hb	9.64	5.31	3.68
			mu (sec)	316	205	141
			sp (sec)	1243	878	355
	Rural	100	efficiency	0.24	0.42	0.90
			ha	34.11	24.06	11.30
			hb	14.92	9.16	5.42
			mu (sec)	529	419	165
			sp (sec)	1564	1004	708
data_sample_2	Urban	43	efficiency	0.21	0.38	1.00
			ha	39.26	21.71	9.98
			hb	16.38	8.07	4.19
			mu (sec)	751	553	100
			sp (sec)	1892	965	477
	Rural	39	efficiency	0.28	0.44	1.00
			ha	23.04	11.86	7.49
			hb	9.28	5.21	3.16
			mu (sec)	316	305	160
			sp (sec)	1423	939	378

Ανάλυση οδηγών – Συγκεντρωτική επίδοση (4/13)

- Χρήση **κινητού τηλεφώνου**
 - Μικρές διαφορές μεταξύ κατηγοριών 1 και 2
 - κατά προσέγγιση ίδια χρήση μεταξύ των οδηγών των κατηγοριών
 - ευαισθησία της DEA σε ακραίες τιμές
- **Υψηλότερος αριθμός** απότομων συμβάντων σε αστικές οδούς ταχείας κυκλοφορίας απ’ ότι σε αστικό οδικό δίκτυο
- Αριθμός συμβάντων $HA = 2 * \text{Αριθμός συμβάντων HB}$

Sample type	Road type	No of drivers	Driving characteristics (/100km)	Efficiency classes		
				Class 1: 0 - 25 % percentile	Class 2: 25 - 75 % percentile	Class 3: 75 - 100 % percentile
data_sample_1	Urban	100	efficiency	0.22	0.36	0.61
			ha	21.49	11.82	8.82
			hb	9.64	5.31	3.68
			mu (sec)	316	205	141
			sp (sec)	1243	878	355
	Rural	100	Efficiency	0.24	0.42	0.90
			ha	34.11	24.06	11.30
			hb	14.92	9.16	5.42
			mu (sec)	529	419	165
			sp (sec)	1564	1004	708
data_sample_2	Urban	43	efficiency	0.21	0.38	1.00
			ha	39.26	21.71	9.98
			hb	16.38	8.07	4.19
			mu (sec)	751	553	100
			sp (sec)	1892	965	477
	Rural	39	efficiency	0.28	0.44	1.00
			ha	23.04	11.86	7.49
			hb	9.28	5.21	3.16
			mu (sec)	316	305	160
			sp (sec)	1423	939	378

Ανάλυση οδηγών – Αποδοτικό επίπεδο (5/13)

	Real level of metrics						Lamdas of peers: Driver No				Efficient level of metrics				
Driver No	distance _{urban}	ha _{urban}	hb _{urban}	speeding _{urban}	mobile _{urban}	Theta	12	34	40	42	distance _{urban}	ha _{urban}	hb _{urban}	speeding _{urban}	mobile _{urban}
1	1868	326	134	21712	9954	0.581	-	0.52	0.14	0.06	3214.3	159.8	77.9	12617.9	5784.7
2	2456	574	85	27049	13974	0.696	-	0.19	0.40	0.20	3526.5	154.8	59.2	18838.2	9732.1
3	1634	709	509	15888	42817	0.391	0.79	-	0.20	-	4182.6	277.0	78.1	6206.9	10434.2
4	2219	233	181	27052	12421	0.637	-	0.23	0.31	0.18	3481.0	148.5	59.1	17244.6	7917.9
5	4223	1088	309	37825	29581	0.613	0.30	1.16	0.47	-	6887.3	417.6	189.5	23192.7	18137.9
6	2773	652	251	25829	25887	0.529	0.60	0.51	0.30	-	5245.3	344.7	128.8	13654.8	13685.4
7	2086	265	149	20880	14036	0.619	-	0.52	0.28	0.04	3371.9	163.9	79.2	12917.2	8683.2
8	1789	460	323	17185	5960	0.656	-	0.75	-	0.01	2728.0	184.3	98.1	11269.8	3908.5
9	1630	184	82	30801	8955	0.528	-	-	0.21	0.22	3085.0	97.2	30.8	14784.6	4731.5
10	808	266	95	9985	6562	0.443	0.11	0.24	0.04	-	1824.6	97.2	42.1	4421.6	2905.8
11	3012	913	152	21604	43562	0.585	0.72	-	0.83	-	5149.6	308.3	88.9	12636.1	23107.5
12	1462	329	92	5074	7781	1.000	1.00	-	-	-	1462.0	329.0	92.0	5074.0	7781.0
*Inputs = [ha _{urban} , hb _{urban} , speeding _{urban} , mobile _{urban}], Output = [distance _{urban}]															

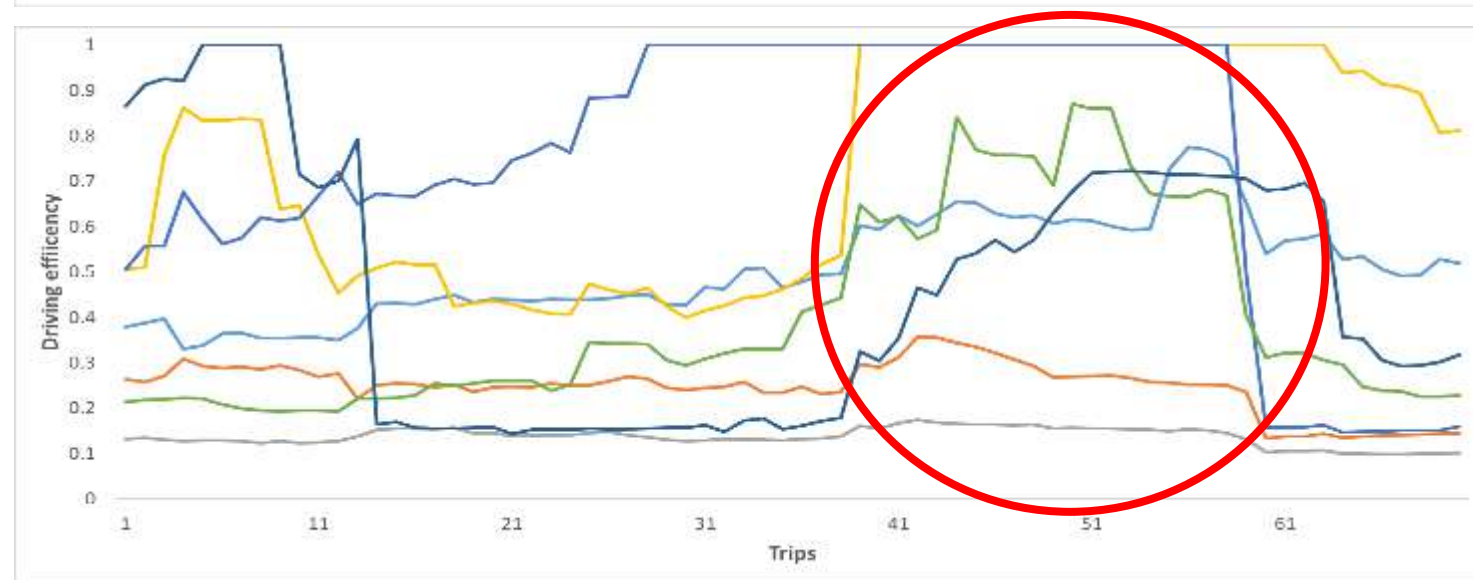
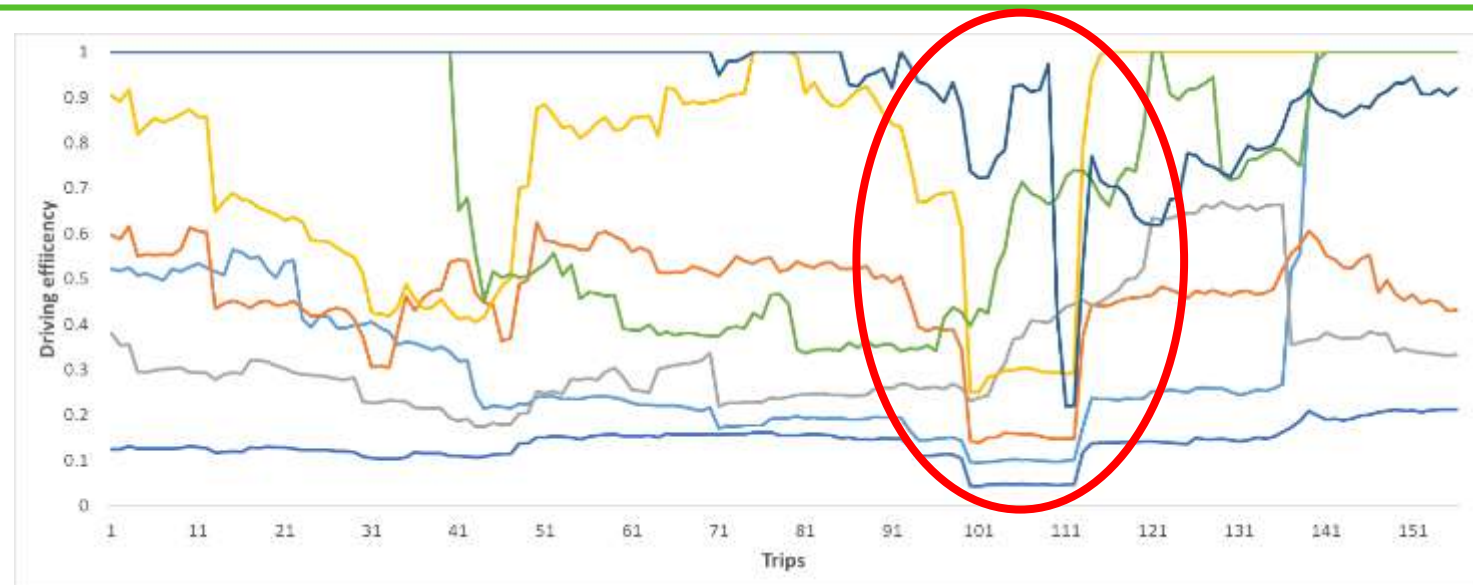
Ανάλυση οδηγών – Χρονική εξέλιξη (6/13)

- Χρονική εξέλιξη

- κυλιόμενο παράθυρο
 - Αστικό οδικό δίκτυο: 75
 - Αστικό οδικό δίκτυο ταχείας κυκλοφορίας: 81

- Οι **λιγότερο αποδοτικοί** οδηγοί έχουν λιγότερο μεταβλητή συμπεριφορά

- Κοινά σημεία τοπικού **ελαχίστου** και **μεγίστου**
 - ύπαρξη ακραίων τιμών
 - η επίδοση υπολογίζεται συγκριτικά



Ανάλυση οδηγών – Χρονική εξέλιξη (7/13)

- Χρονοσειρά

- Στασιμότητα

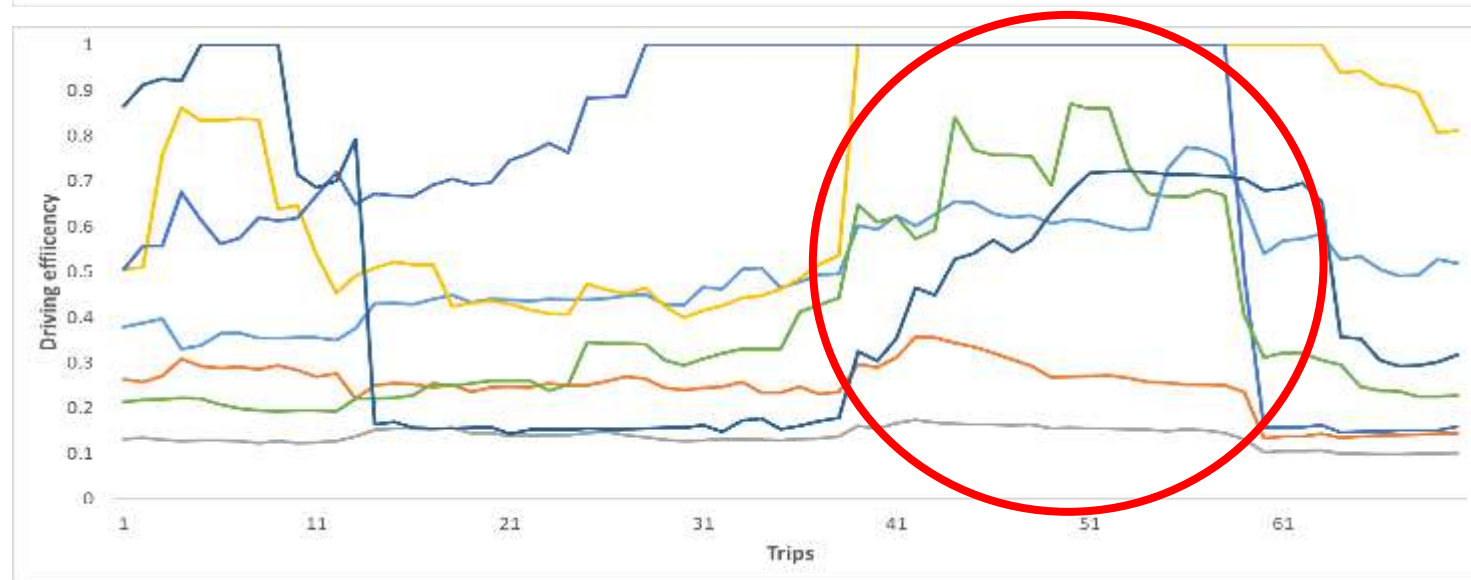
- κανένας στάσιμος οδηγός σε αστικό
- ασήμαντος αριθμός στάσιμων οδηγών σε αστικό οδικό δίκτυο ταχείας κυκλοφορίας

- Τάση

- ο μέσος όρος είναι κατά προσέγγιση ίδιος στους 2 τύπους οδών
- υψηλότερο εύρος στο αστικό οδικό δίκτυο ταχείας κυκλοφορίας

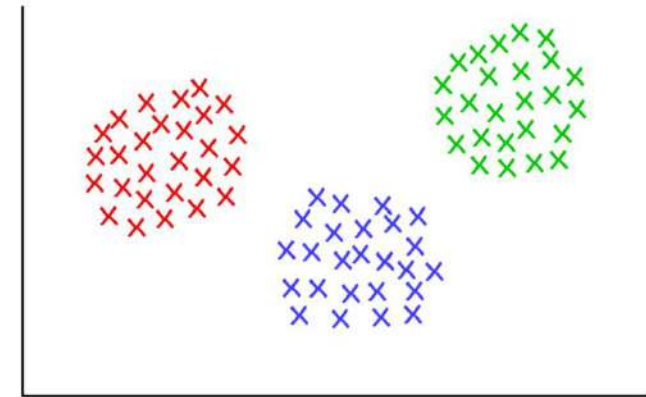
- Μεταβλητότητα

- ο μέσος όρος είναι κατά προσέγγιση ίδιος στους 2 τύπους οδών
- υψηλότερο εύρος στο αστικό οδικό δίκτυο ταχείας κυκλοφορίας



Ανάλυση οδηγών – Συσταδοποίηση (8/13)

- **Τρεις συστάδες** για τα 2 δείγματα
 - συστάδα 1: τυπικός οδηγός
 - συστάδα 2: ασταθής οδηγός
 - συστάδα 3: προσεκτικός οδηγός



Ανάλυση οδηγών – Συσταδοποίηση (9/13)

- **Τυπικοί** οδηγοί:
 - μεγαλύτερο αριθμό οδηγών
 - χαμηλή συνολική επίδοση
 - πολύ χαμηλή θετική τάση
 - μέτρια προς υψηλή μεταβλητότητα
 - χαμηλή συχνότητα ατυχημάτων

Sample type	Road type	Cluster	Trend (*10 ⁻³)	Volatility	Efficiency	Accidents/ 10 years of driving experience
data_sample_1	Urban	1 (typical)	very low positive	medium - high	low	low - medium
		2 (unstable)	medium positive	medium - high	medium	low
		3 (cautious)	medium negative	low - medium	medium - high	low
	Rural	1 (typical)	low positive	medium	low	low - medium
		2 (unstable)	high negative	high	medium - high	medium - high
		3 (cautious)	high positive	medium - high	high	low
data_sample_2	Urban	1 (typical)	very low positive	medium	low	low
		2 (unstable)	low - medium positive	medium	low	high
		3 (cautious)	medium negative	low	high	low
	Rural	1 (typical)	barely no trend	medium - high	low	low
		2 (unstable)	low negative	medium	low	high
		3 (cautious)	high positive	medium - high	high	low

Ανάλυση οδηγών – Συσταδοποίηση (10/13)

- **Ασταθείς** οδηγοί:
 - επίδοση
 - μέτρια προς υψηλή συνολική επίδοση για το data_sample_1
 - χαμηλή συνολική για το data_sample_2
 - τάση
 - μέτρια θετική για το data_sample_1
 - μέτρια αρνητική για το data_sample_2
- μέτρια προς υψηλή μεταβλητότητα
 - υψηλή για το data_sample_1
 - μέτρια για το data_sample_2
- μέτρια προς υψηλή συχνότητα ατυχημάτων

Sample type	Road type	Cluster	Trend (*10 ⁻³)	Volatility	Efficiency	Accidents/ 10 years of driving experience
data_sample_1	Urban	1 (typical)	very low positive	medium - high	low	low - medium
		2 (unstable)	medium positive	medium - high	medium	low
		3 (cautious)	medium negative	low - medium	medium - high	low
	Rural	1 (typical)	low positive	medium	low	low - medium
		2 (unstable)	high negative	high	medium - high	medium - high
		3 (cautious)	high positive	medium - high	high	low
data_sample_2	Urban	1 (typical)	very low positive	medium	low	low
		2 (unstable)	low - medium positive	medium	low	high
		3 (cautious)	medium negative	low	high	low
	Rural	1 (typical)	barely no trend	medium - high	low	low
		2 (unstable)	low negative	medium	low	high
		3 (cautious)	high positive	medium - high	high	low

Ανάλυση οδηγών – Συσταδοποίηση (11/13)

- Προσεκτικοί οδηγοί:
 - υψηλή συνολική επίδοση
 - τάση
 - μέτρια αρνητική για αστικό οδικό δίκτυο
 - υψηλά θετική για αστικό οδικό δίκτυο ταχείας κυκλοφορίας
- μέτρια μεταβλητότητα
 - μέτρια προς χαμηλή για το αστικό οδικό δίκτυο
 - μέτρια προς υψηλή για το αστικό οδικό δίκτυο ταχείας κυκλοφορίας
- χαμηλή συχνότητα ατυχημάτων

Sample type	Road type	Cluster	Trend (*10 ⁻³)	Volatility	Efficiency	Accidents/ 10 years of driving experience
data_sample_1	Urban	1 (typical)	very low positive	medium - high	low	low - medium
		2 (unstable)	medium positive	medium - high	medium	low
		3 (cautious)	medium negative	low - medium	medium - high	low
	Rural	1 (typical)	low positive	medium	low	low - medium
		2 (unstable)	high negative	high	medium - high	medium - high
		3 (cautious)	high positive	medium - high	high	low
data_sample_2	Urban	1 (typical)	very low positive	medium	low	low
		2 (unstable)	low - medium positive	medium	low	high
		3 (cautious)	medium negative	low	high	low
	Rural	1 (typical)	barely no trend	medium - high	low	low
		2 (unstable)	low negative	medium	low	high
		3 (cautious)	high positive	medium - high	high	low

Ανάλυση οδηγών – Συσταδοποίηση (12/13)

- Οι τιμές των χαρακτηριστικών **μειώνονται** κατά τη μετάβαση σε συστάδα **υψηλότερης επίδοσης**
- **Τυπικοί** οδηγοί
 - διπλάσιες τιμές χαρακτηριστικών οδήγησης σε αστικό απ’ ότι σε αστικό οδικό δίκτυο ταχείας κυκλοφορίας
- **Προσεκτικοί** οδηγοί
 - σημαντικά χαμηλότερο επίπεδο για όλα τα χαρακτηριστικά οδήγησης

Sample type	Road type	No of drivers	Driving characteristics	Cluster 1 (typical)	Cluster 2 (unstable)	Cluster 3 (cautious)
data_sample_1	Urban	100	efficiency	0.33	0.61	0.81
			ha	26.75	17.20	6.89
			hb	10.05	7.30	3.44
			mu	499	165	60
			sp	1095	619	1240
	Rural	100	efficiency	0.36	0.69	0.88
			ha	14.97	6.36	9.65
			hb	7.59	3.09	3.61
			mu	234	285	86
			sp	988	923	347
data_sample_2	Urban	43	efficiency	0.37	0.33	1.00
			ha	22.05	41.26	14.13
			hb	8.10	15.39	5.28
			mu	653	481	82
			sp	1349	1140	436
	Rural	39	efficiency	0.37	0.39	1.00
			ha	12.35	19.24	5.74
			hb	6.66	6.84	2.46
			mu	316	380	149
			sp	1125	1149	415

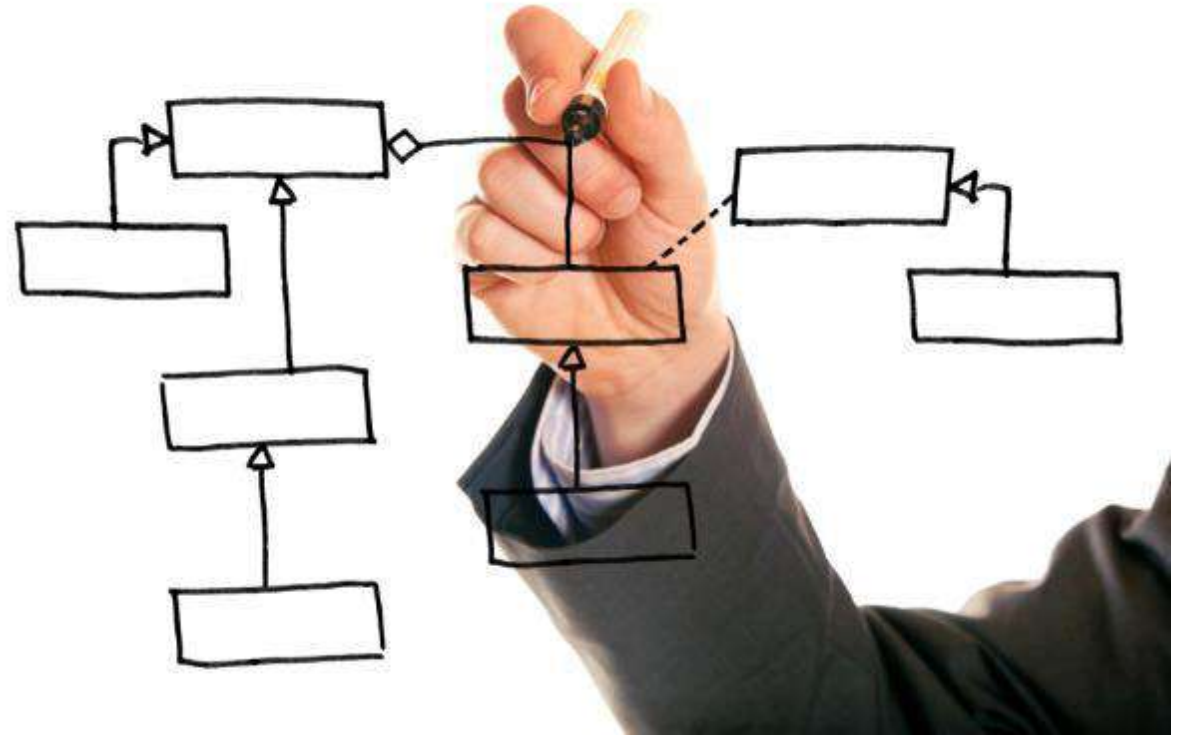
Ανάλυση οδηγών – Συσταδοποίηση (13/13)

- **Ασταθείς** οδηγοί του data_sample_2
 - υψηλότερος αριθμός HA και HB
 - υψηλότερη χρήση κινητού τηλεφώνου και υπέρβαση ορίου ταχύτητας σε αστικό οδικό δίκτυο ταχείας κυκλοφορίας
- Οδηγοί μεταβλητής συμπεριφοράς με **κακή** οδηγική συμπεριφορά => οδηγοί με την **υψηλότερη** πιθανότητα ατυχήματος

Sample type	Road type	No of drivers	Driving characteristics	Cluster 1 (typical)	Cluster 2 (unstable)	Cluster 3 (cautious)
data_sample_1	Urban	100	efficiency	0.33	0.61	0.81
			ha	26.75	17.20	6.89
			hb	10.05	7.30	3.44
			mu	499	165	60
			sp	1095	619	1240
	Rural	100	efficiency	0.36	0.69	0.88
			ha	14.97	6.36	9.65
			hb	7.59	3.09	3.61
			mu	234	285	86
			sp	988	923	347
data_sample_2	Urban	43	efficiency	0.37	0.33	1.00
			ha	22.05	41.26	14.13
			hb	8.10	15.39	5.28
			mu	653	481	82
			sp	1349	1140	436
	Rural	39	efficiency	0.37	0.39	1.00
			ha	12.35	19.24	5.74
			hb	6.66	6.84	2.46
			mu	316	380	149
			sp	1125	1149	415

Συμπεράσματα (1/5)

- **Καινοτόμος μεθοδολογική προσέγγιση** για τη συγκριτική αξιολόγηση της οδηγικής επίδοσης και για την εκτίμηση του αποδοτικού επιπέδου των χαρακτηριστικών οδήγησης του κάθε οδηγού
- Η **ενσωμάτωση** του αλγορίθμου **convex hull** στην DEA απέδωσε σημαντικά βελτιωμένα αποτελέσματα από τις υπόλοιπες μεθοδολογίες



Συμπεράσματα (2/5)

- Η απαιτούμενη απόσταση **δειγματοληψίας** προσδιορίζεται και είναι διαφορετική για κάθε:
 - τύπο οδού
 - χαρακτηριστικό οδήγησης
 - επίπεδο επιθετικότητας/ επικινδυνότητας/ απόσπασης προσοχής
- Η απαιτούμενη απόσταση δειγματοληψίας δεν καθορίζεται από **ένα χαρακτηριστικό**
- Οι **πιο** επιθετικοί/ επικίνδυνοι οδηγοί χρειάζονται **λιγότερη** παρατήρηση στο αστικό οδικό δίκτυο ταχείας κυκλοφορίας και στο υπεραστικό δίκτυο



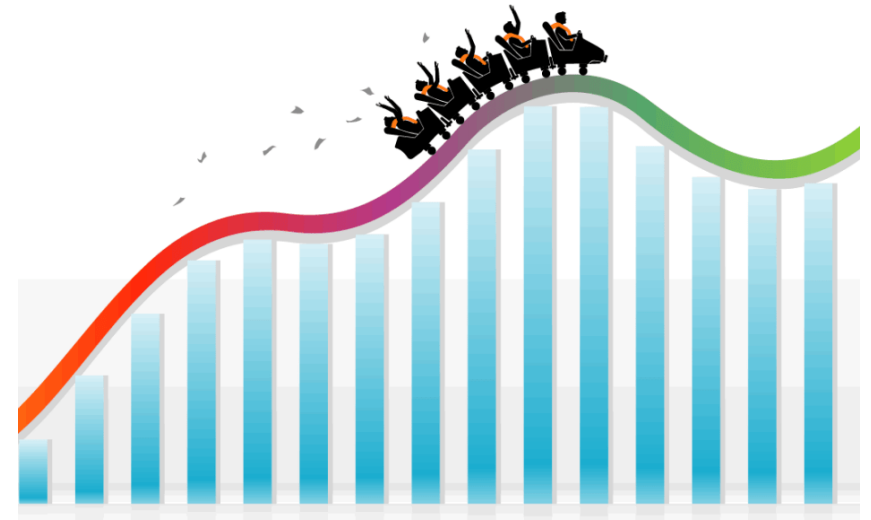
Συμπεράσματα (3/5)

- **Ασήμαντες διαφορές** στη χρήση κινητού τηλεφώνου μεταξύ οδηγών μέτριας και χαμηλής επίδοσης
- Ο αριθμός των απότομων **επιταχύνσεων** είναι **διπλάσιος** από τον αριθμό των απότομων **επιβραδύνσεων**
- Η **αλλαγή** κατηγορίας επίδοσης επηρεάζεται από **διαφορετικά χαρακτηριστικά** οδήγησης σε αστικό οδικό δίκτυο και αστικό οδικό δίκτυο ταχείας κυκλοφορίας



Συμπεράσματα (4/5)

- Η μέση **μεταβλητότητα** είναι κατά προσέγγιση **ίδια** στους 2 τύπους οδών
- Η μέση **τάση** είναι κατά προσέγγιση **ίδια** στους 2 τύπους οδών
- Η **στασιμότητα** είναι παρόμοια για όλους τους οδηγούς και τους τύπους οδών



Συμπεράσματα (5/5)

- Παρατηρούνται **3 συστάδες** οδηγών:
α) τυπικοί, β) ασταθείς και γ) προσεκτικοί οδηγοί
- Οι οδηγοί πρέπει να παρακολουθούνται **συνεχώς** και να **επαναξιολογούνται** προκειμένου να παρατηρηθούν τυχόν μεταβολές
- **Πληροφόρηση** για το ιστορικό **ατυχημάτων** επηρεάζει μόνο την σύσταση της ομάδας των **ασταθών** οδηγών



Συνεισφορά

- Αναπτύχθηκε **μεθοδολογικό πλαίσιο** για την αξιολόγηση της οδηγικής επίδοσης σε επίπεδο διαδρομής και οδηγού βασιζόμενο σε τεχνικές της επιστήμης των δεδομένων
- **Ποσοτικοποιήθηκε** η απαιτούμενη απόσταση δειγματοληψίας που θα πρέπει να συλλέγεται για την αξιολόγηση της οδηγικής συμπεριφοράς από άποψη ασφαλείας
- Παρέχονται πληροφορίες για τα κυριότερα χαρακτηριστικά των **οδηγικών προφίλ** που παρατηρούνται
- Χρησιμοποιείται ένα **καινοτόμο** σύστημα **συλλογής δεδομένων** οδήγησης



Αξιοποίηση

- Παροχή προσωποποιημένης και όχι μόνο **πληροφόρησης** στους οδηγούς για
 - την συνολική τους επίδοση και την εξέλιξή της
 - όταν μια μη αποδοτική διαδρομή λαμβάνει χώρα
 - οδηγικά χαρακτηριστικά που πρέπει να βελτιωθούν
 - σε κάθε τύπο οδού
- **Μείωση** του κινδύνου στην οδήγηση
- Καθορισμός πολιτικής τιμολόγησης για εταιρείες **ασφάλισης**
 - χρέωση ασφαλιστρων με βάση την οδηγική συμπεριφορά



Μελλοντικές προκλήσεις (1/2)

- Αξιοποίηση **μεγαλύτερου** οδηγικού **δείγματος**
 - μελέτη υπεραστικού οδικού δικτύου
 - πιο αντιπροσωπευτική εκτίμηση της οδηγικής επίδοσης ως προς την ασφάλεια
 - αντιμετώπιση ευαισθησίας της DEA στις ακραίες τιμές
 - εξέλιξη των χαρακτηριστικών οδήγησης
 - σχέση της επιθετικότητας των οδηγών και της απαιτούμενης απόστασης δειγματοληψίας
 - γνωστό το ιστορικό ατυχημάτων
 - διερεύνηση για περαιτέρω μείωση του χρόνου υπολογισμού
- Τύπος **προσέγγισης**
 - μακροσκοπική
 - μεταβολές στη συμπεριφορά σε μεγάλα χρονικά διαστήματα
 - μικροσκοπική
 - μελέτη των χωρο-χρονικών χαρακτηριστικών κάθε διαδρομής
 - συνδυασμός
- **Πρόβλεψη** της οδηγικής επίδοσης/ κατηγορίας/ συστάδας
 - προτάσεις για να επιτευχθεί το αποδοτικό επίπεδο των χαρακτηριστικών οδήγησης



Μελλοντικές προκλήσεις (2/2)

- **Περιορισμοί DEA** – μηδενικό άθροισμα DEA inputs
- **Αύξηση** του αριθμού των χαρακτηριστικών οδήγησης
 - απόσταση από προπορευόμενο όχημα
 - αλλαγή λωρίδας
 - κίνηση των ματιών
 - υπνηλία
- **Συλλογή δεδομένων** μέσω:
 - κάμερας
 - συσκευών παρατήρησης των κινήσεων των ματιών
 - Radars
 - LiDARs
 - συστημάτων καταγραφής εντός του οχήματος (OBD)
- Μελέτη της **δυναμικής εξέλιξης** της οδηγικής επίδοσης προκειμένου να ποσοτικοποιηθεί
 - ταχύτητα μεταβολής του οδηγικού προφίλ
 - το μέγεθος της μεταβολής του οδηγικού προφίλ



*Thank
you*



Συγκριτική ανάλυση της οδηγικής επίδοσης με χρήση τεχνικών της επιστήμης των δεδομένων εφαρμοζόμενων σε δεδομένα κινητών τηλεφώνων μεγάλης κλίμακας



Δημήτρης Τσελέντης

Πολιτικός – Συγκοινωνιολόγος Μηχανικός
Υποψήφιος Διδάκτωρ – Ερευνητής

www.nrso.ntua.gr/dtsel/
dtsel@central.ntua.gr