CAS: Crime Anticipation System

Predictive Policing in Amsterdam

« waakzaam en dienstbaar »

POLITIE

Introduction

- **Dick Willems**
  - Background in Mathematical Psychology (University of Nijmegen)
  - Statistician at Universities of Nijmegen and Maastricht
  - Datamining consultant for commercial businesses
  - Joined Amsterdam Police Department in 2012

Agenda

- **Police Datamining in Amsterdam**
- **Predictive policing: Crime Anticipation System**
  - Origin
  - Method
  - Future

Police Datamining in Amsterdam

- **The Netherlands are divided into 10 regional units, of which Amsterdam is one.**
- **The Dutch Police puts great effort into reducing High Impact Crimes (domestic burglary, mugging and robbery)**
- **Some numbers concerning High Impact Crimes in Amsterdam:**
  - Domestic burglary: 8.257 incidents in 2013
  - Mugging: 2.358 incidents in 2013
  - Robbery: 276 incidents in 2013
- **One of the tactics the Amsterdam Police uses is to intelligently allocate manpower where and when it matters most, using data mining methods.**

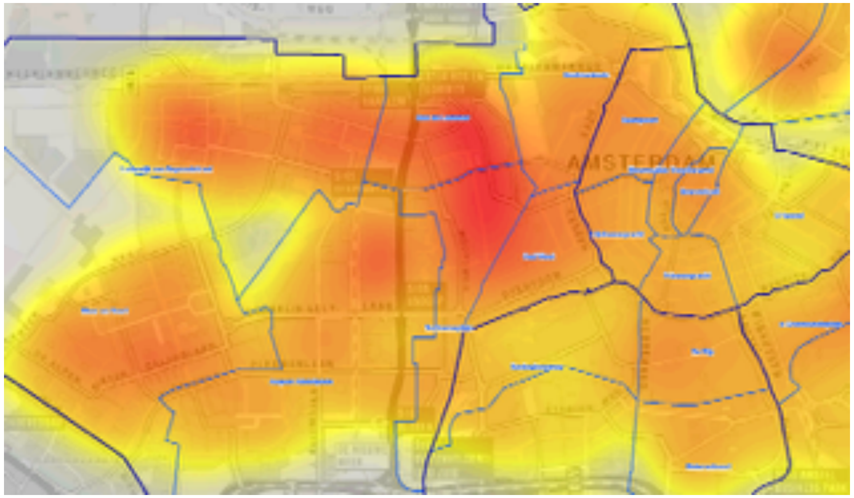
Police Datamining in Amsterdam

- **The Amsterdam Police Department has invested in datamining for over 12 years.**
  - 2 dataminers in service
  - Availability of data mining software
  - Availability of dedicated server

- **Team datamining works on:**
  - Predictive policing
  - Extraction of useful information from texts in police reports (domestic violence, discrimination, human trafficking, identification of potentially dangerous “einzelgangers”)
  - Uncovering criminal networks
  - Sporadic issues involving large amounts of data

Predictive Policing: Crime Anticipation System
Origins

- Planning of manpower used to take place using “gut feeling” and ad hoc analyses.
- Police analysts were capable of making “hot spot” maps: plots of incident locations modified by applying a Gaussian filter

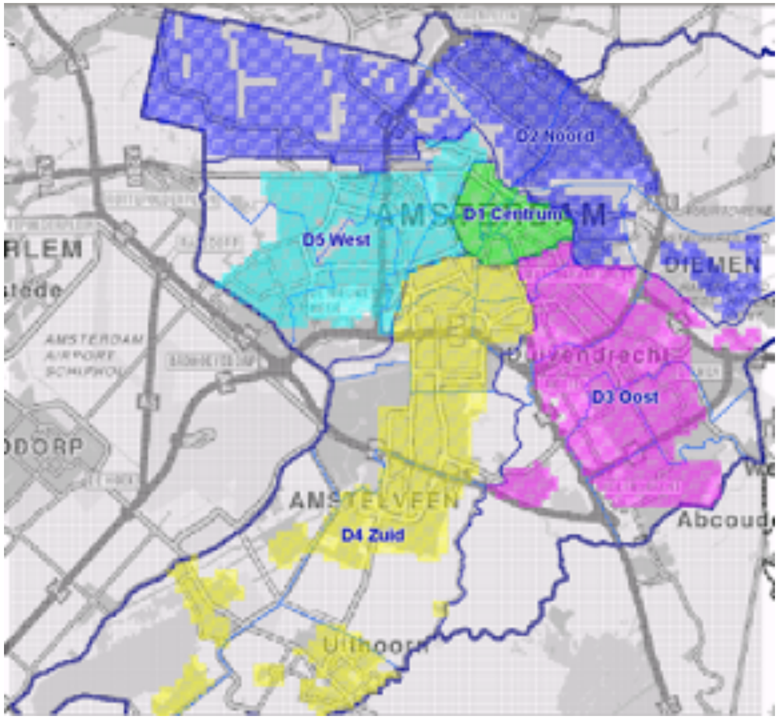


- A grid divides the Amsterdam map into 125m by 125m squares.

- There will be more events in some squares than in others.
- Determine characteristics of the squares from the database.
- Calculate the probability of an event in a square based on its characteristics.

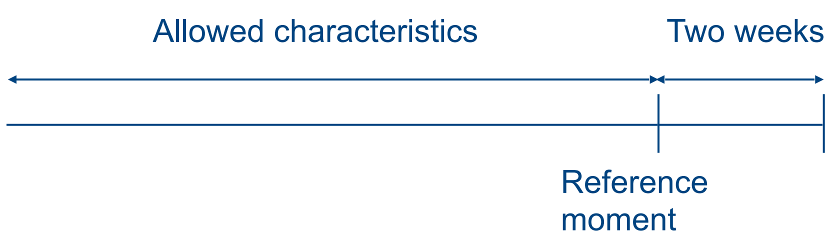
- Knowing the locations, determine when (what day, what time) the risk on an event is greatest.

- First selection: exclude squares that are “empty” (pastures, open water, et cetera)
- This leaves 11.500 relevant squares of 196x196 = 38.416 possible ones.
- Collect data from the remaining squares for three years (reference moments every two weeks).
- Every square has 78 data points.
- Total dataset has 78x11500 = 897.000 data points.



- **For each reference moment and square, a number of characteristics are computed:**
  - Location specific characteristics
  - Crime history
  - Which crimes took place within the two weeks following the reference moment.

- **For prediction, it's important that only those characteristics are recorded that could have been known at the reference moment.**



Predictive Policing: Crime Anticipation System
Method

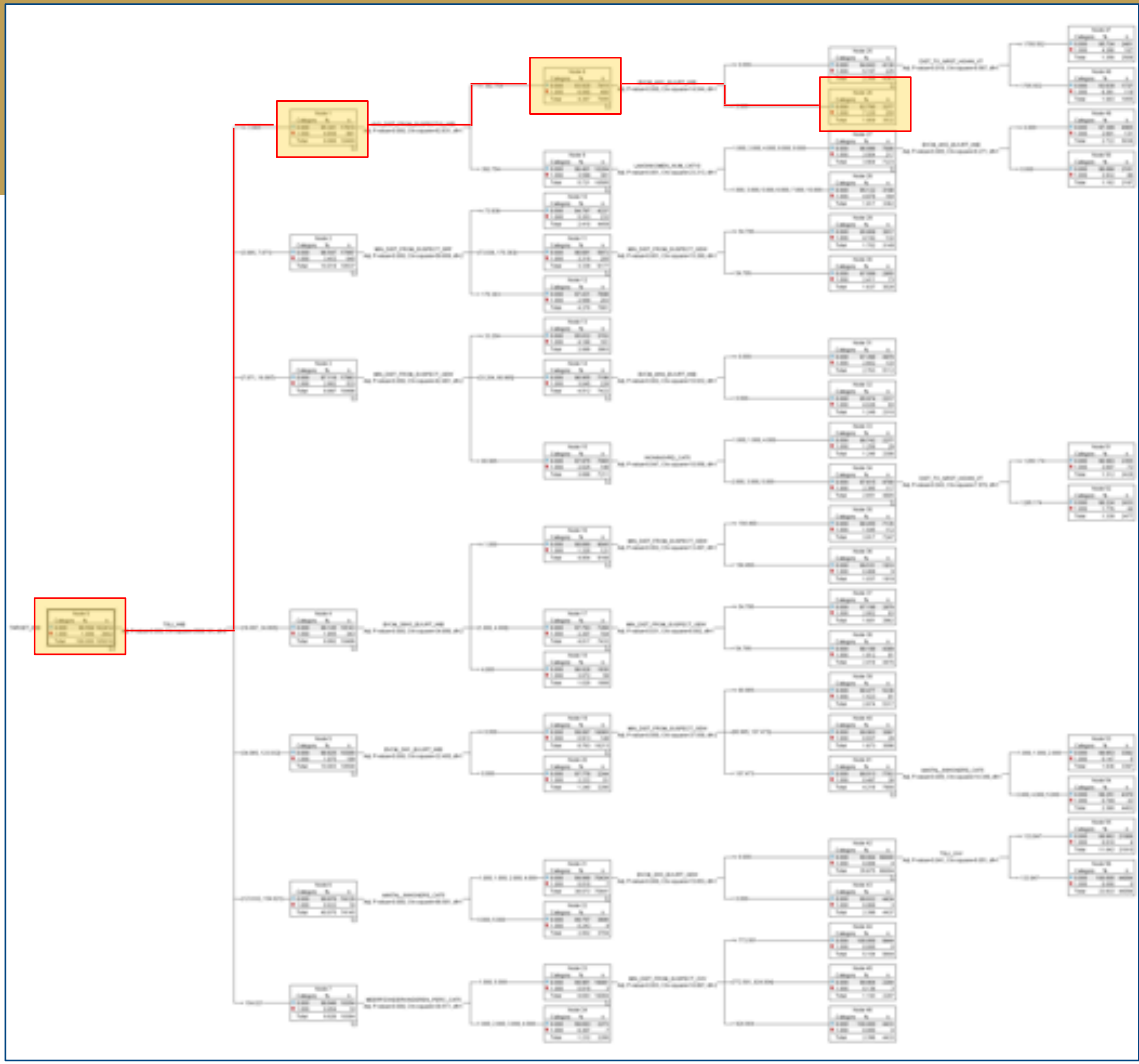
- Static, location specific characteristics
  - Information from the central bureau of statistics: demografics en socioeconomical characteristics
  - Number and kind of companies (bars, coffeeshops, banks, etc)
  - Distance to the closest known offender (mugger, robber, burglar etc)
  - Mean distance to the 10 closest known offenders
  - Distance to the nearest highway exit

- Crime history
  - Number of burglaries, robberies etc in several different time periods (relative to reference moment)
    - Last two weeks, two weeks before that etc
    - Last four weeks, four weeks before that etc
    - Last half year
  - Same for the neighbouring squares
  - Linear trend in square (number of crimes increasing, decreasing, stable?)
  - Season

Predictive Policing: Crime Anticipation System
Method

- To link the characteristics known at the reference moment to what happens afterwards, an artificial neural network was applied.
- Result: a model that can assign risk-scores to squares based on current knowledge.

Simplified model



Burglary

No rule applied: all squares have the same probability of burglary in the next two weeks (1.4%).

If the last burglary in the square was less than 3 months ago, this rises to 4.7%

If additionally the mean distance to the 10 closest known burglars is less than 400m , then this rises to 6.1%

If a burglary has taken place in the last four weeks, this becomes 7.2%

Simplified model

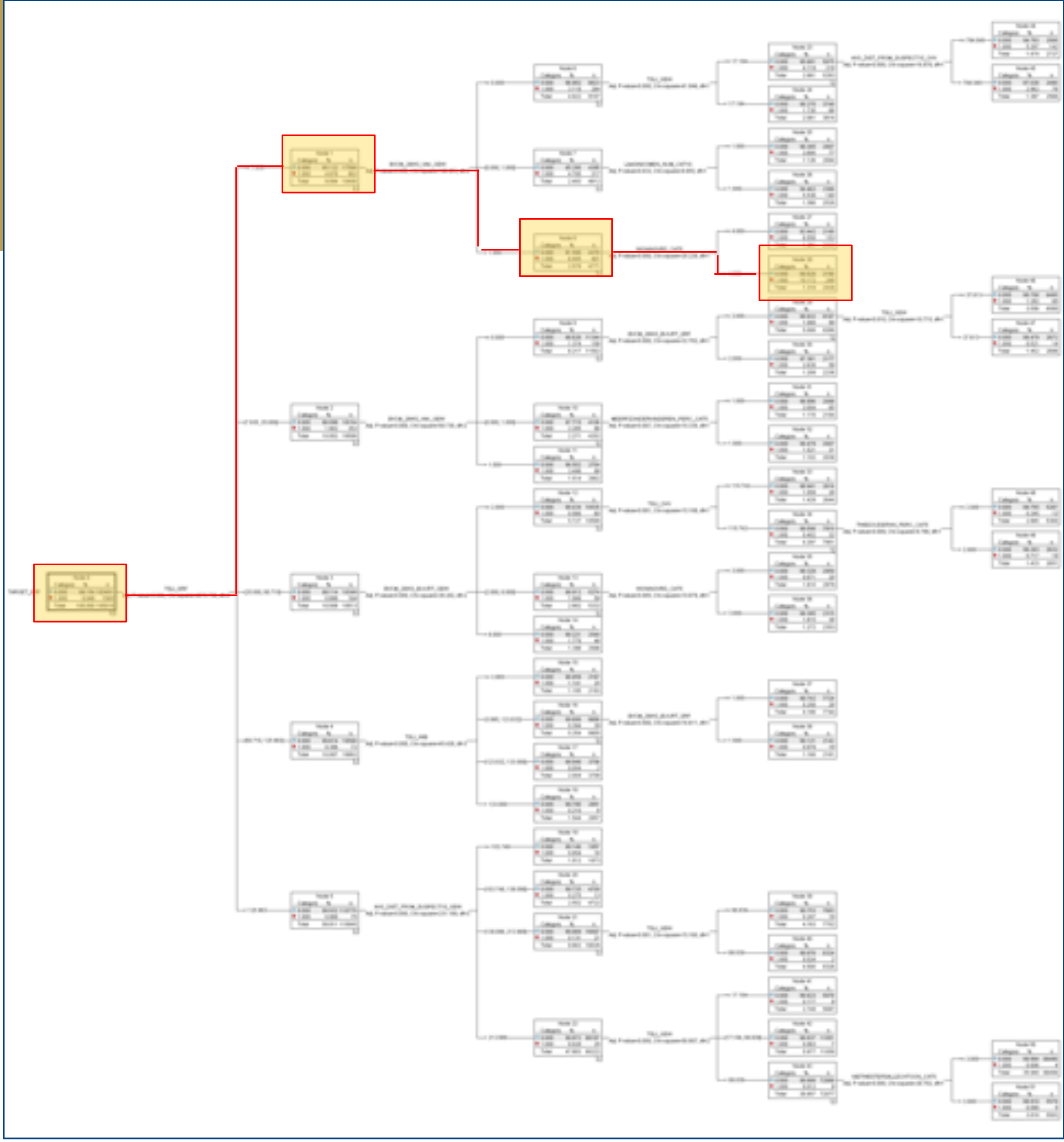
Mugging

No rule applied: alls squares have the same probablility of a mugging in the next two weeks (0.8%).

If a mugging has occurred in the square less than 8 months ago, this increases to 4.9%

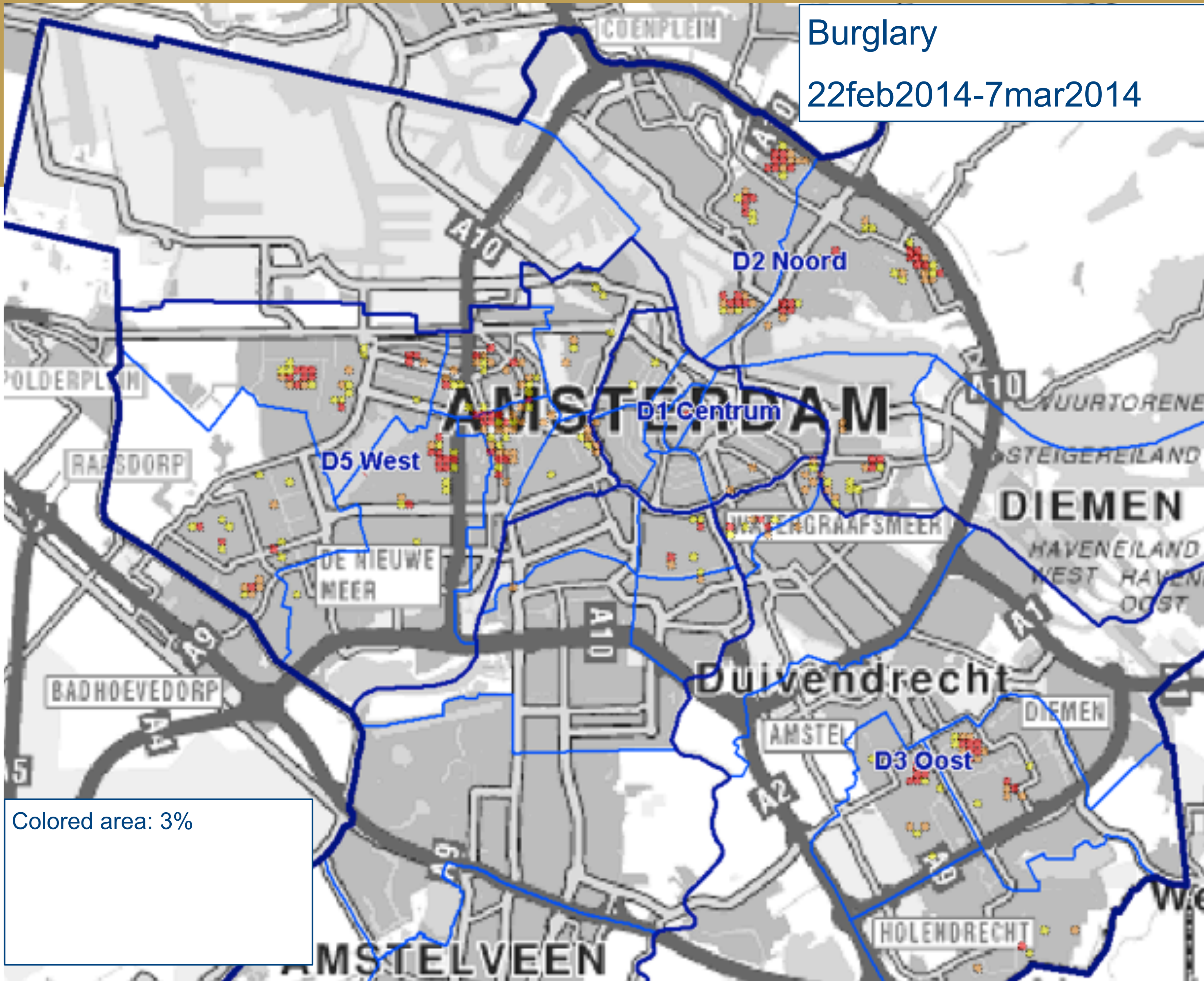
If in the last half year two or more violent incidents have taken place in the square, this rises to 8.4%

If there are many houses in the square, the probability increases to 10.2%.



Burglary

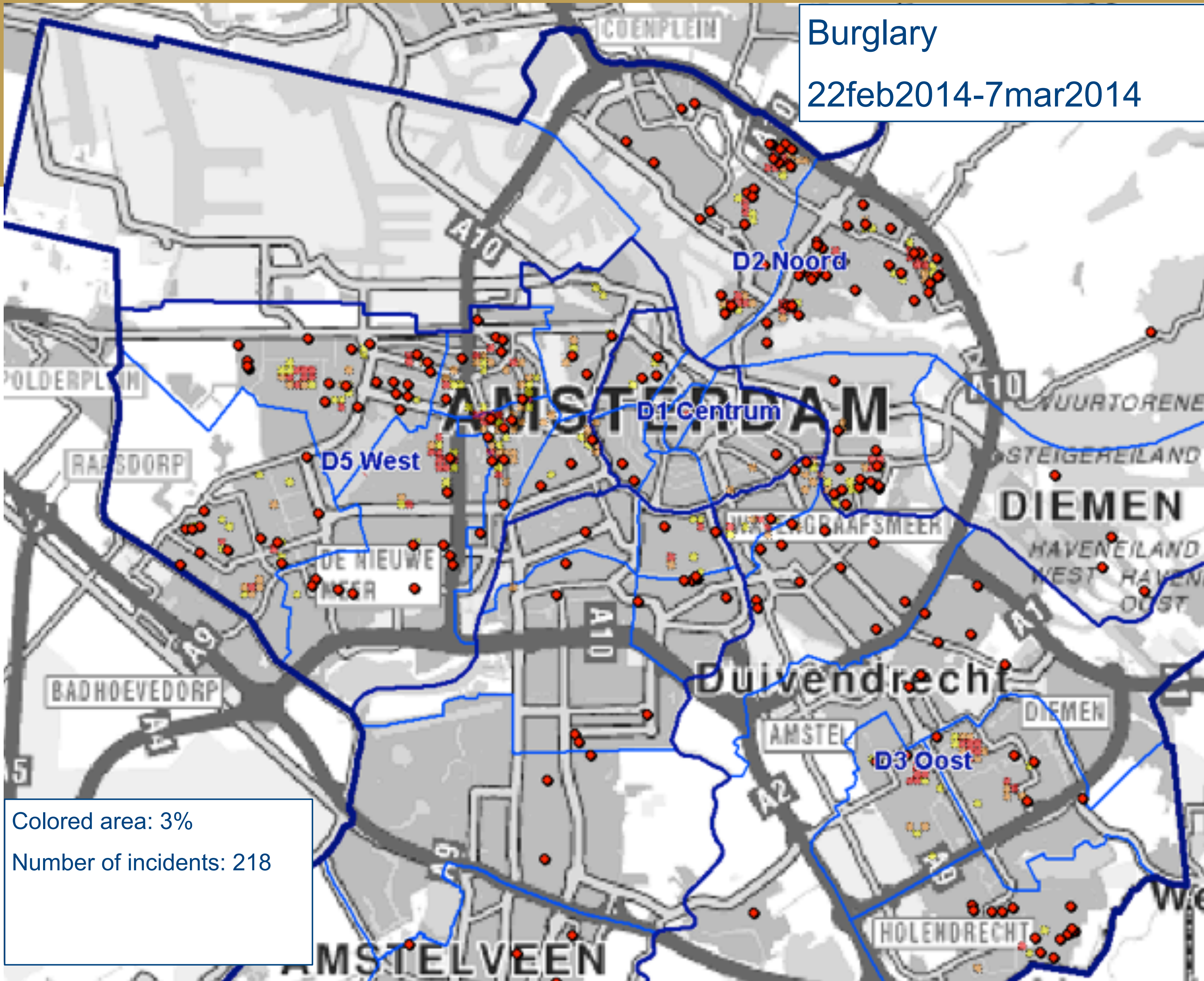
22feb2014-7mar2014



Colored area: 3%

Burglary

22feb2014-7mar2014

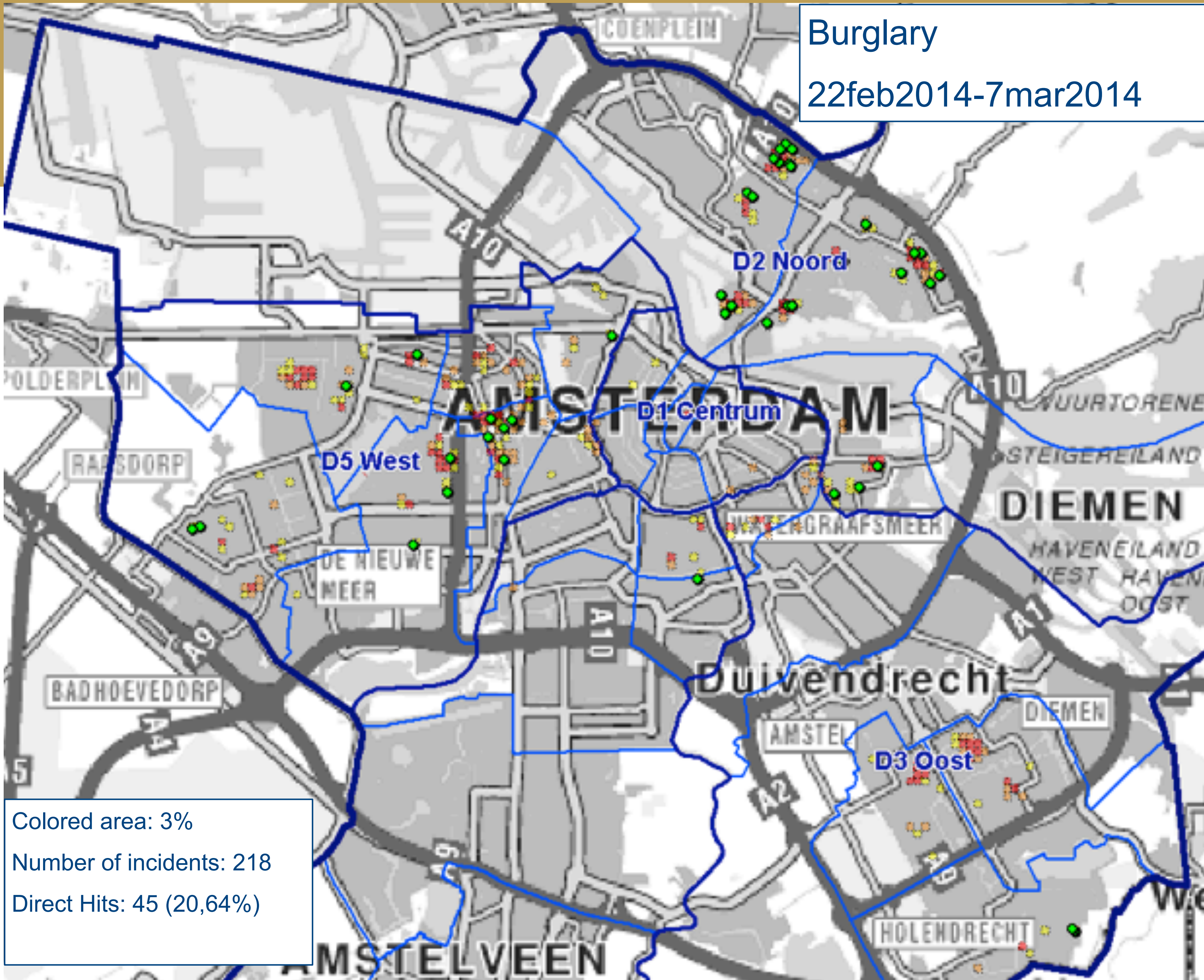


Colored area: 3%

Number of incidents: 218

Burglary

22feb2014-7mar2014



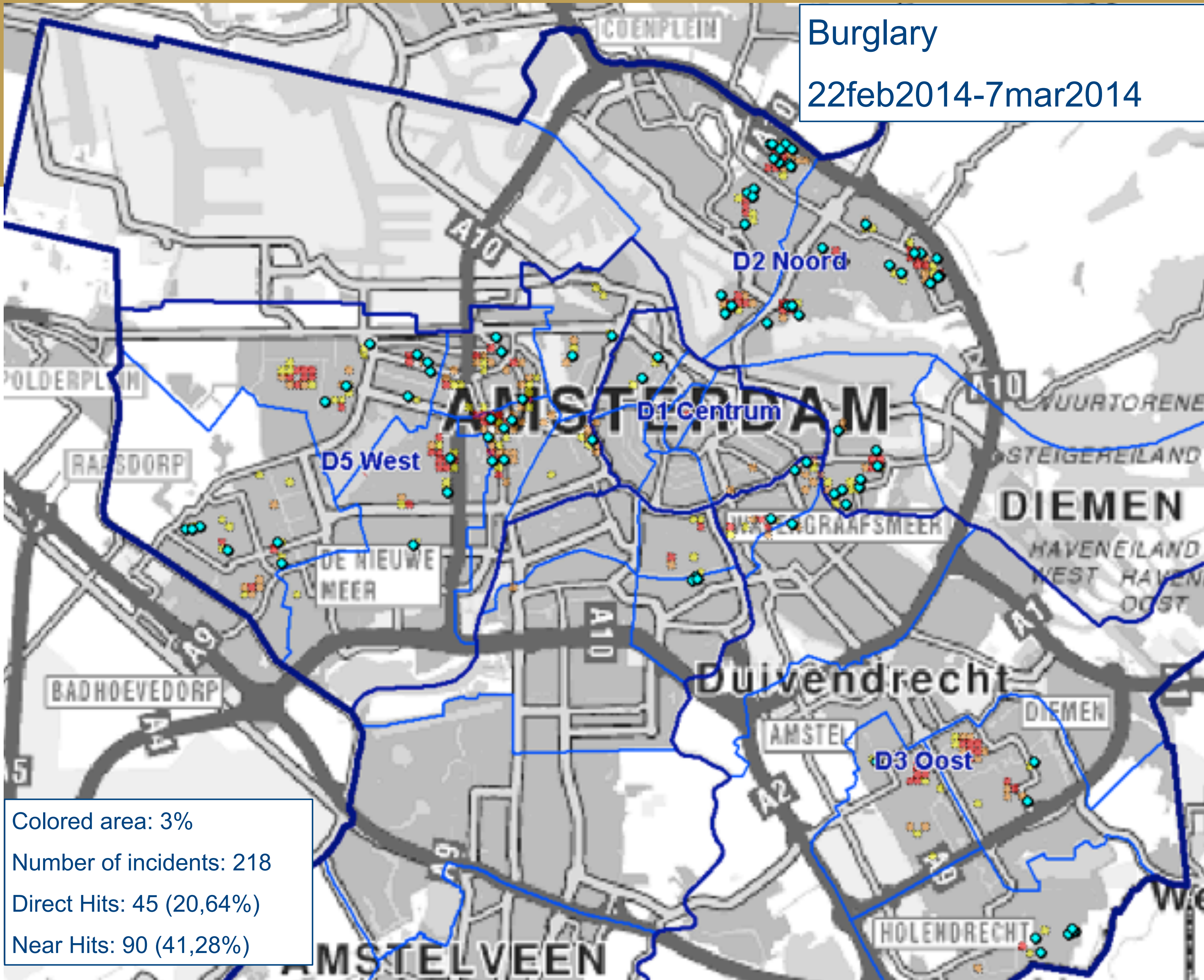
Colored area: 3%

Number of incidents: 218

Direct Hits: 45 (20,64%)

Burglary

22feb2014-7mar2014



Colored area: 3%

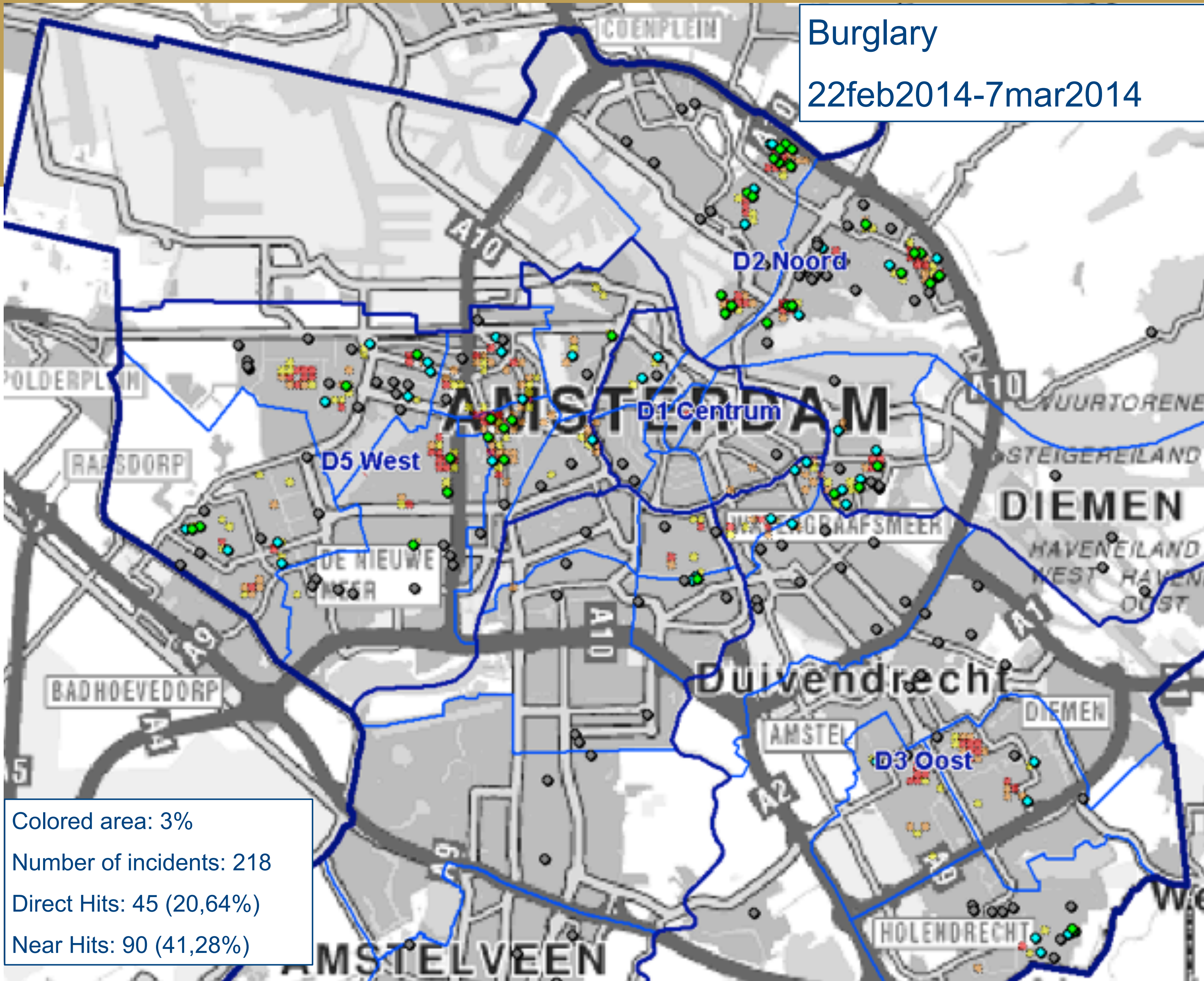
Number of incidents: 218

Direct Hits: 45 (20,64%)

Near Hits: 90 (41,28%)

Burglary

22feb2014-7mar2014

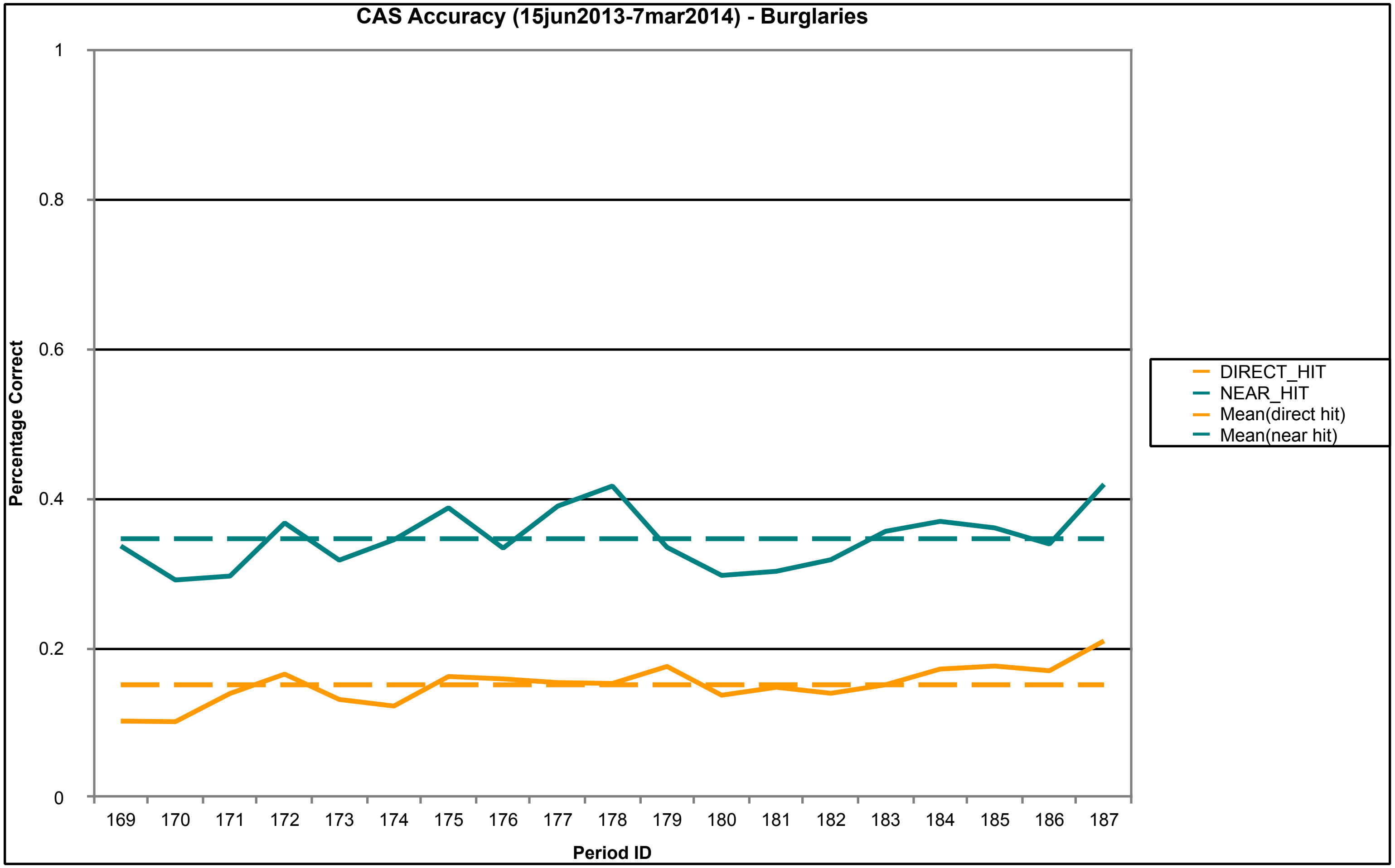


Colored area: 3%

Number of incidents: 218

Direct Hits: 45 (20,64%)

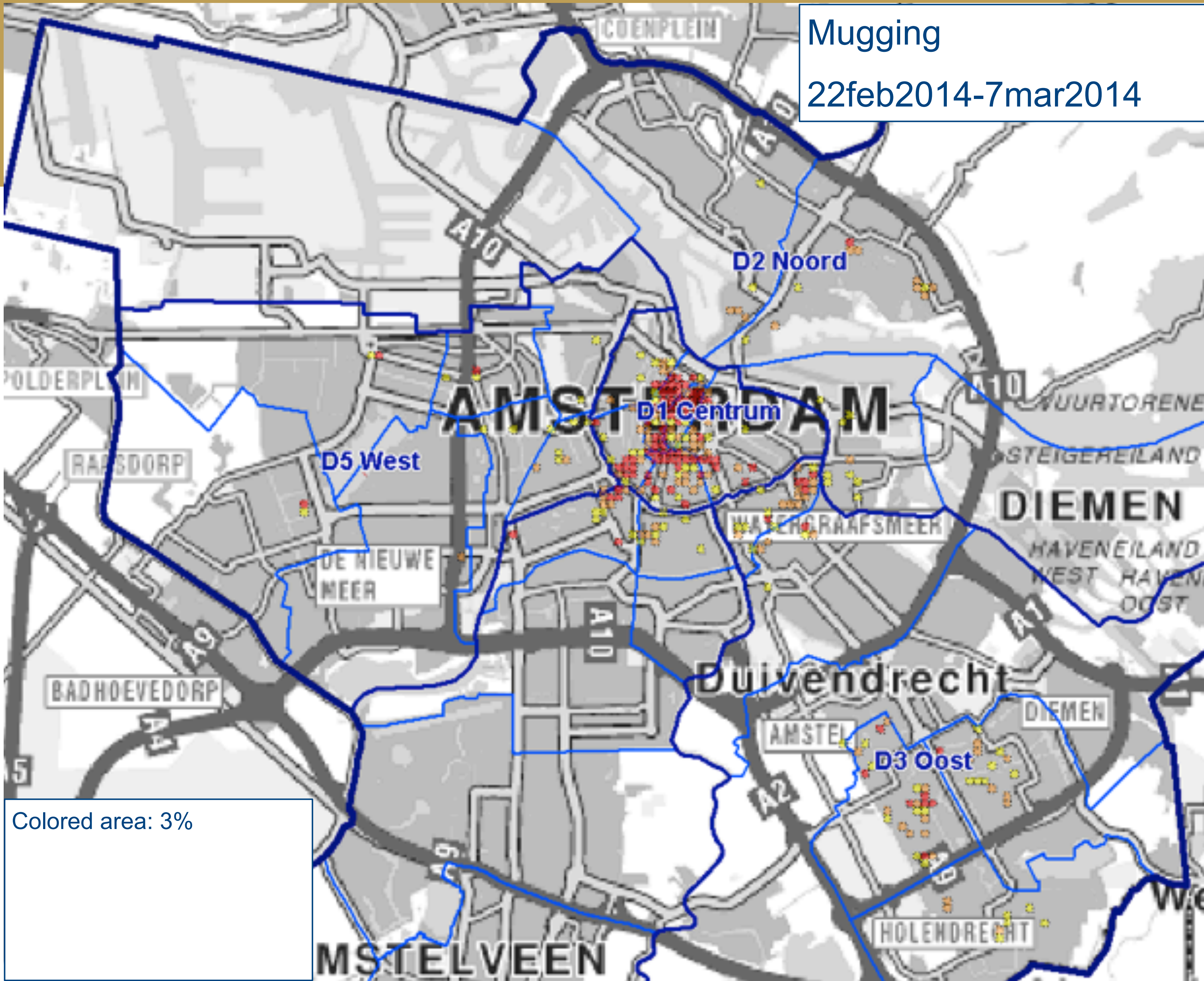
Near Hits: 90 (41,28%)



CAS Accuracy (15jun2013-7mar2014) - Burglaries
Percentage Correct
1
0.8
0.6
0.4
0.2
0
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
Period ID
DIRECT_HIT
NEAR_HIT
Mean(direct hit)
Mean(near hit)

Mugging

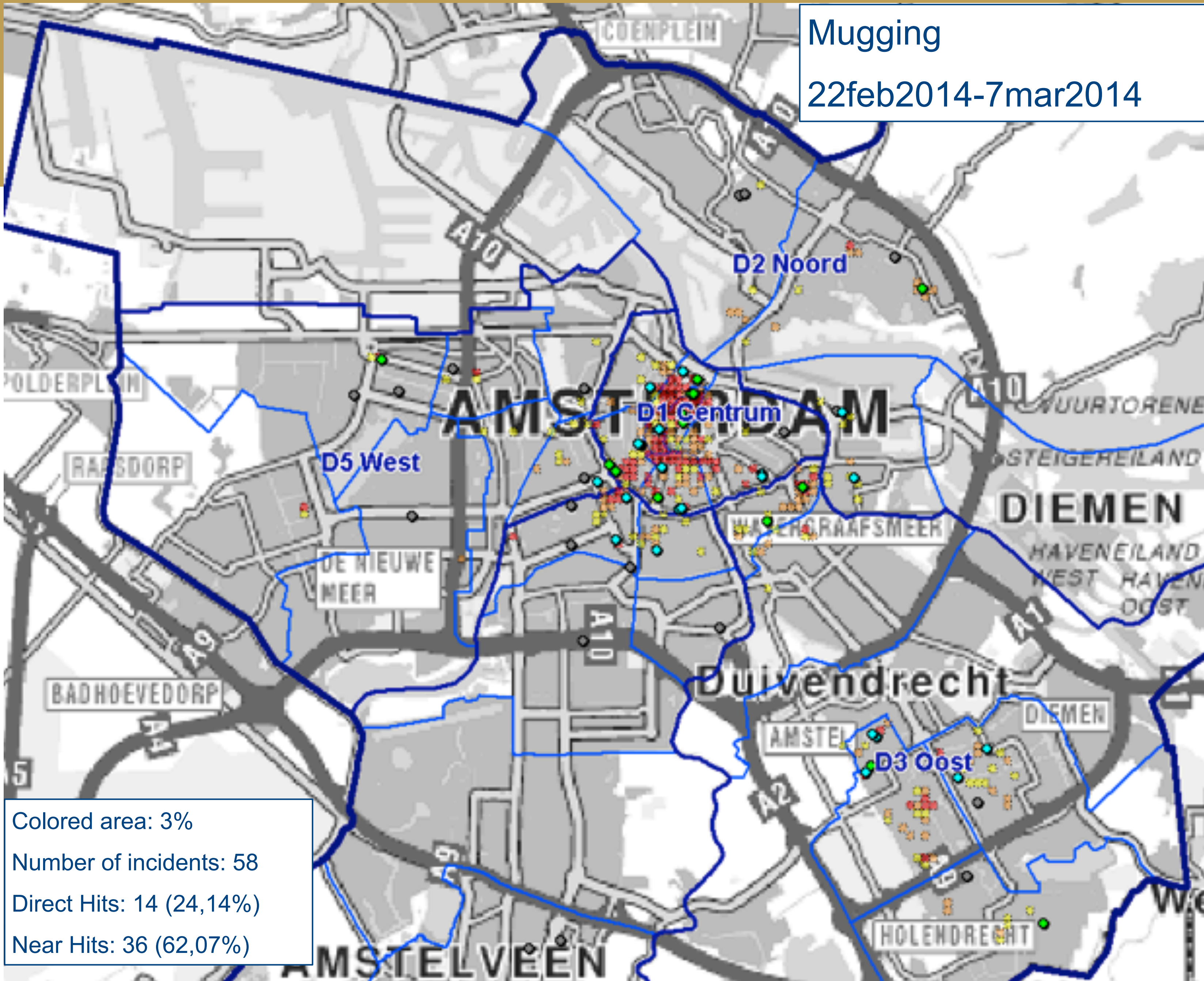
22feb2014-7mar2014



Colored area: 3%

Mugging

22feb2014-7mar2014

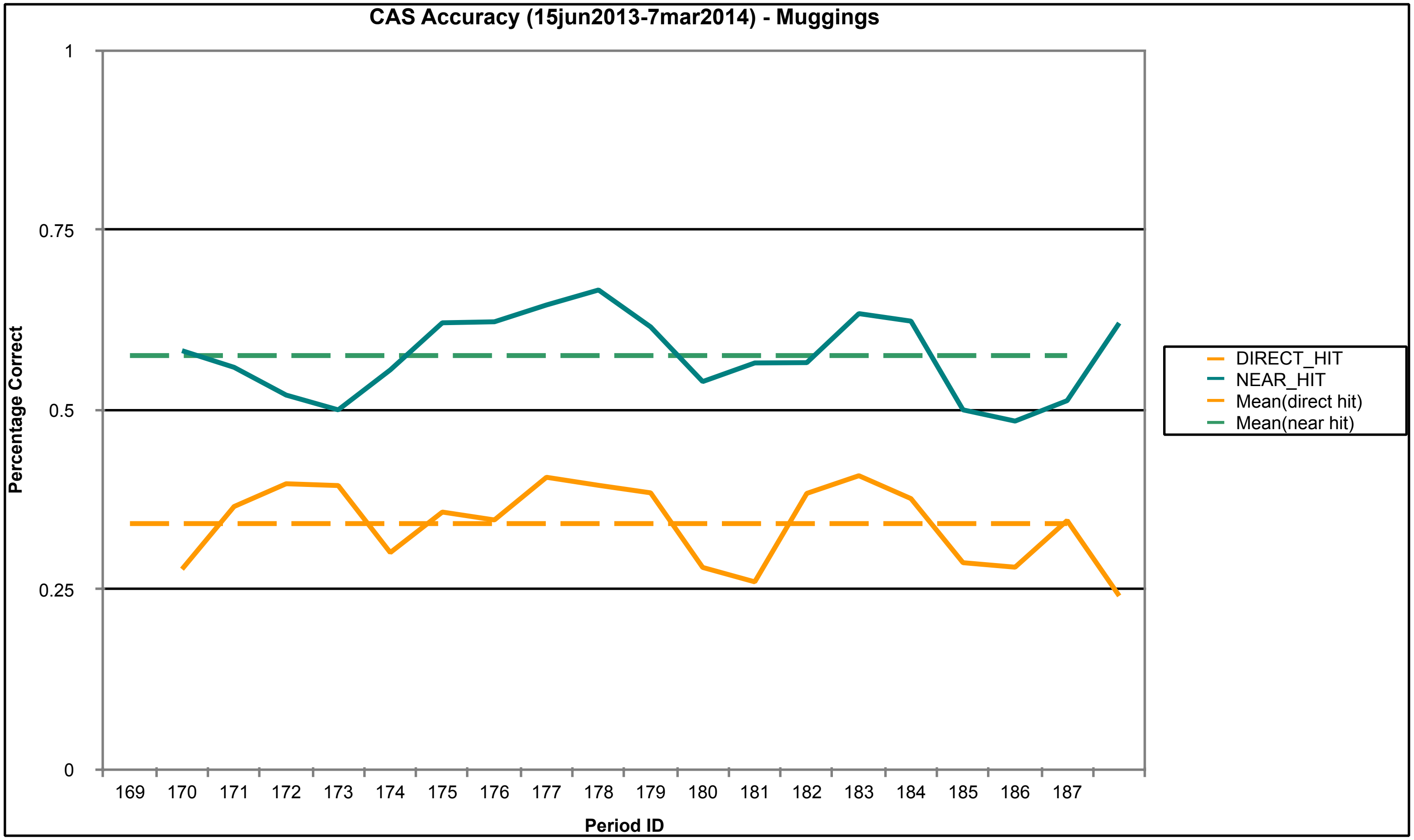


Colored area: 3%

Number of incidents: 58

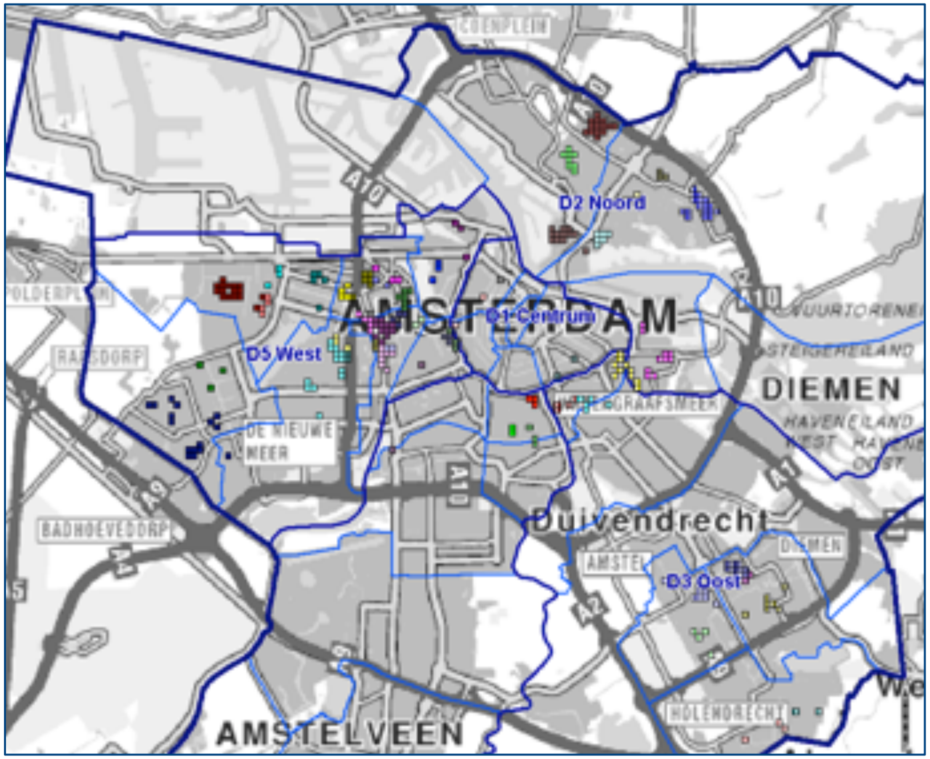
Direct Hits: 14 (24,14%)

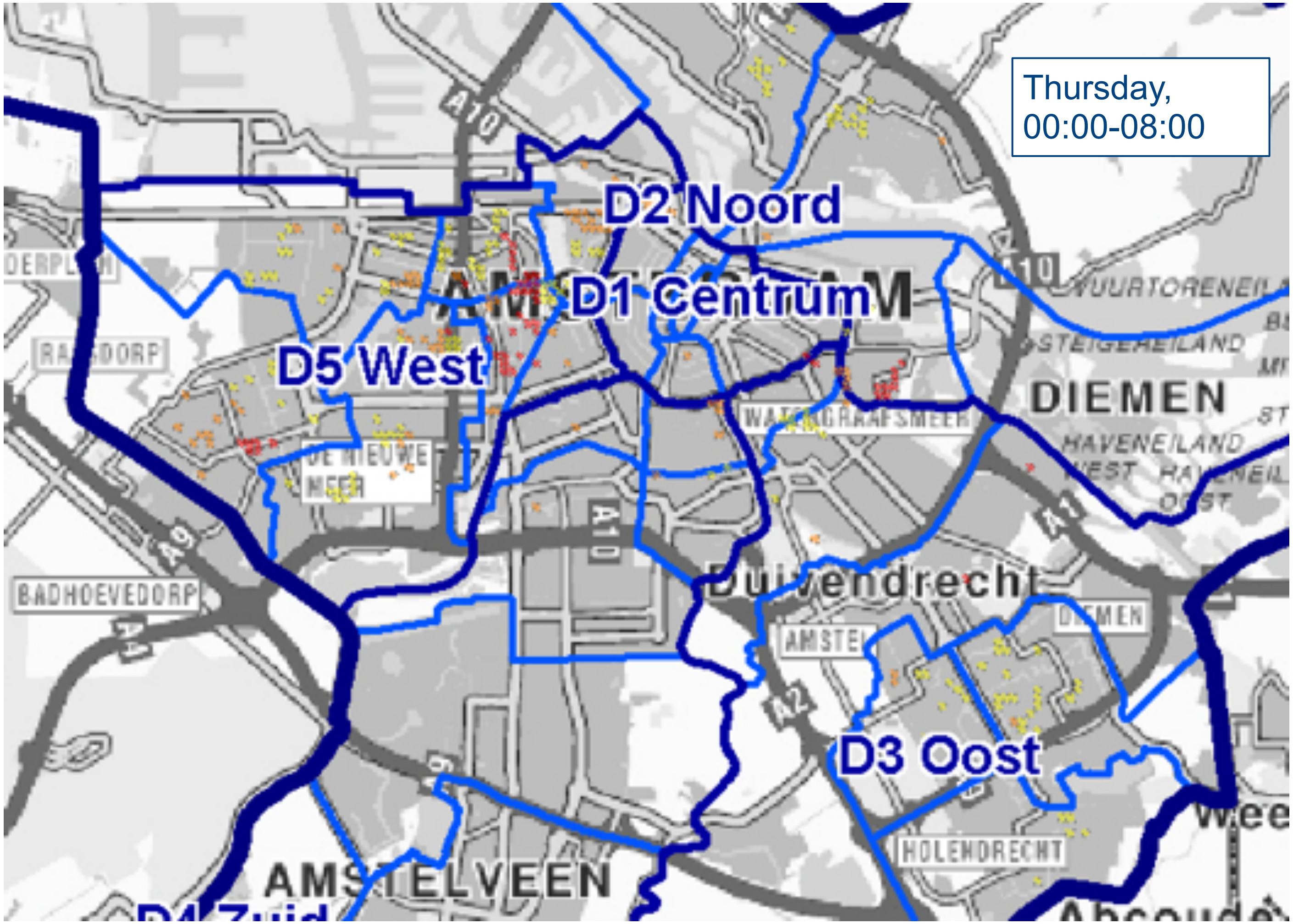
Near Hits: 36 (62,07%)

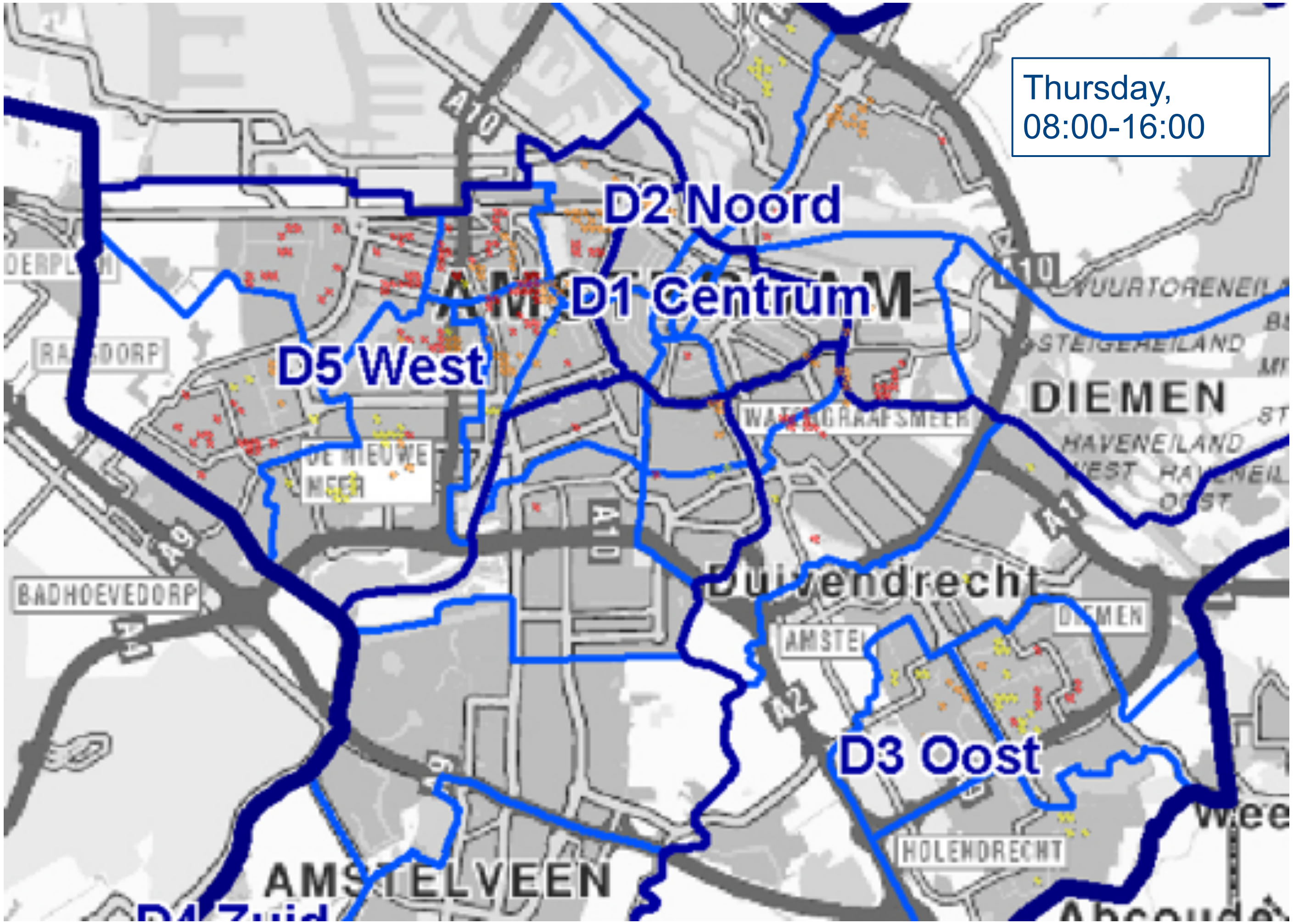


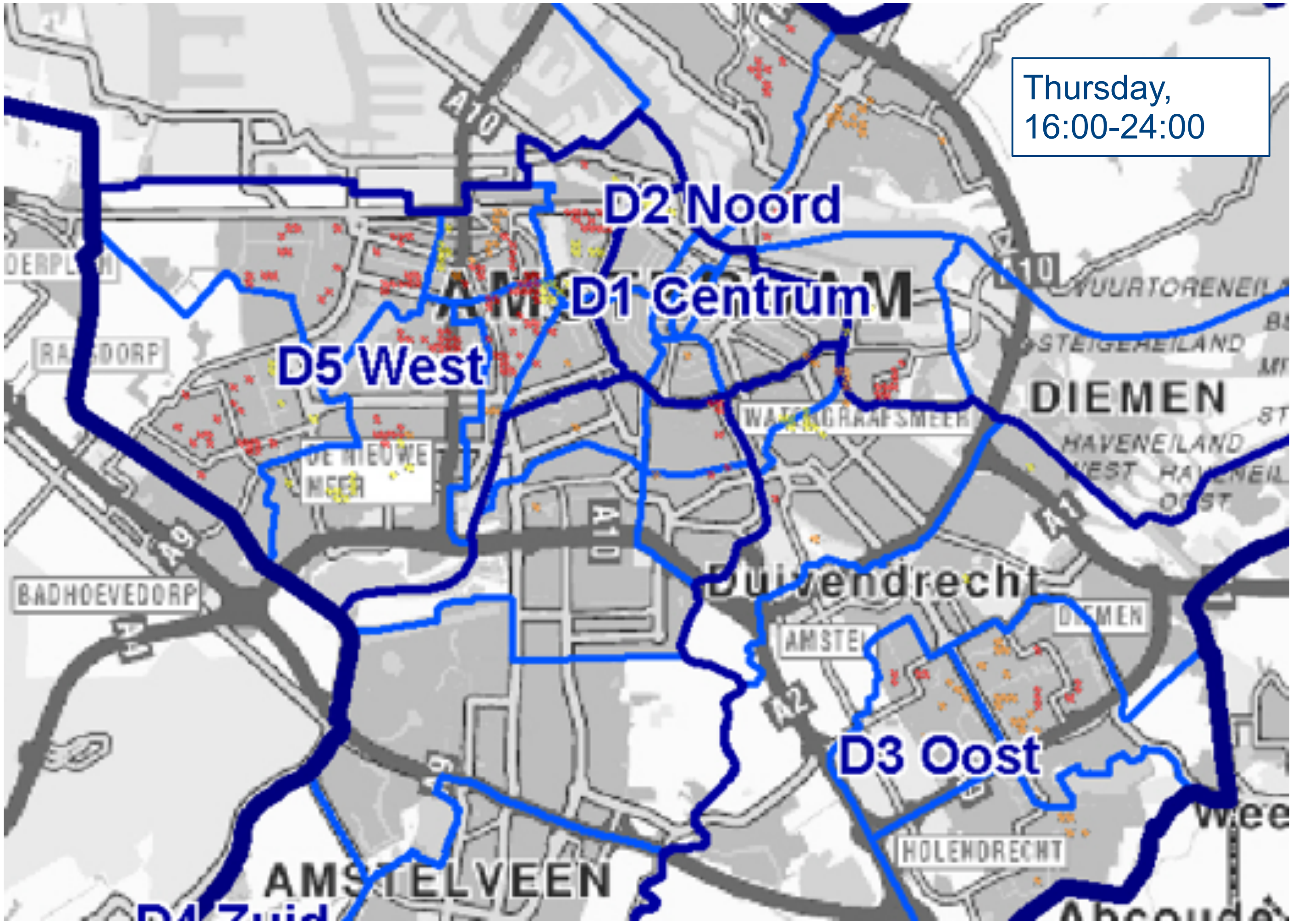
CAS Accuracy (15jun2013-7mar2014) - Muggings
Percentage Correct
1
0.75
0.5
0.25
0
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
Period ID
DIRECT_HIT
NEAR_HIT
Mean(direct hit)
Mean(near hit)

- **High-risk times are determined after high-risk locations are identified**
  - Locations are geographically clustered using a Kohonen algorithm
  - Weekday of incident is recorded
  - Incident times are categorized to correspond to police shifts
    - 00:00-08:00
    - 08:00-16:00
    - 16:00-24:00
- **The characteristics above are used in a similar classification algorithm.**




Thursday,
00:00-08:00
D2 Noord
D1 Centrum
D5 West
D3 Oost
D4 Zuid
DIEMEN
Duivendrecht
AMSTELVEEN
BADHOEVEDORP
HOLENDRECHT
AMSTEL
DE NIEUWE MEER
A10
A9
A4
A2
A1


Thursday,
08:00-16:00
D2 Noord
D1 Centrum
D5 West
D3 Oost
D4 Zuid
DIEMEN
AMSTELVEEN
Duivendrecht
BADHOEVEDORP
HOLENDRECHT
A10
A9
A4
A2
A1


Thursday,
16:00-24:00
D2 Noord
D1 Centrum
D5 West
D3 Oost
D4 Zuid
AMSTERDAM
DIEMEN
Duivendrecht
AMSTELVEEN
WATERGRAAFSMEER
BADHOEVEDORP
DE NIEUWE MEER
AMSTEL
HOLENDRECHT
HAVENEILAND WEST
HAVENEILAND OOST
STEIGEREILAND
VUURTORENEILAND
A10
A9
A4
A2
A1

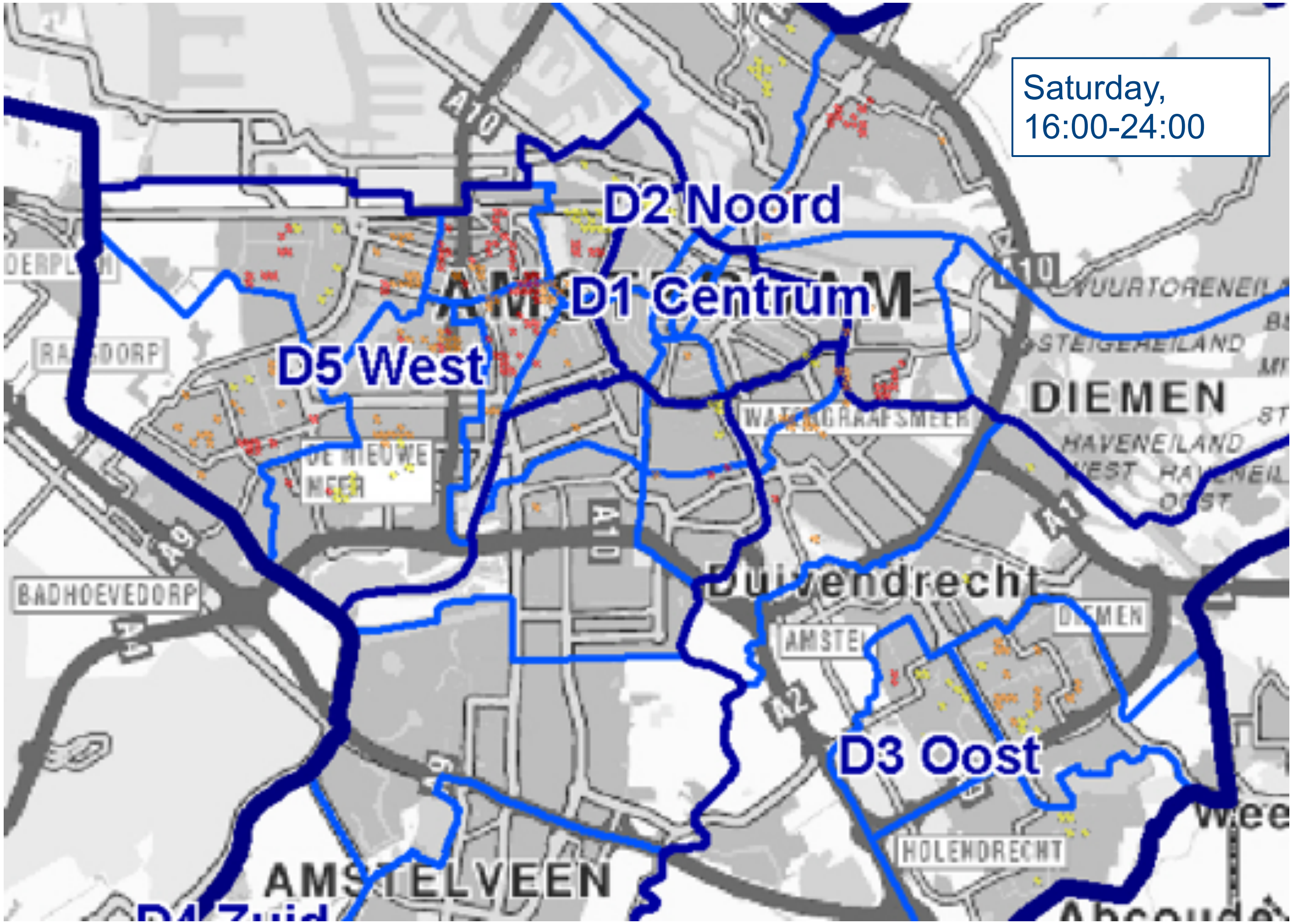

Friday,
00:00-08:00
D2 Noord
D1 Centrum
D5 West
D3 Oost
D4 Zuid
DIEMEN
Duivendrecht
AMSTELVEEN
WATERGRAAFSMEER
BADHOEVEDORP
HOLENDRECHT
AMSTEL
A10
A9
A4
A2
A1


Friday,
08:00-16:00
D2 Noord
D1 Centrum
D5 West
D3 Oost
D4 Zuid
DIEMEN
Duivendrecht
AMSTELVEEN
A10
A9
A4
A2
A1
WATERGRAAFSMEER
BADHOEVEDORP
HOLENDRECHT
AMSTEL
DIEMEN

Friday,
16:00-24:00
D2 Noord
D1 Centrum
D5 West
DIEMEN
Duivendrecht
D3 Oost
AMSTELVEEN
BADHOEVEDORP
HOLENDRECHT
AMSTEL
A10


Saturday,
00:00-08:00
D2 Noord
D1 Centrum
D5 West
D3 Oost
D4 Zuid
DIEMEN
Duivendrecht
AMSTELVEEN
BADHOEVEDORP
HOLENDRECHT
WATERGRAAFSMEER
DE NIEUWE MEER
A10


Saturday,
08:00-16:00
D2 Noord
D1 Centrum
D5 West
D3 Oost
D4 Zuid
DIEMEN
Duivendrecht
AMSTELVEEN
BADHOEVEDORP
HOLENDRECHT
AMSTEL
A10
A9
A4
A2
A1


Saturday,
16:00-24:00
D2 Noord
D1 Centrum
D5 West
D3 Oost
DIEMEN
Duivendrecht
AMSTELVEEN
BADHOEVEDORP
HOLENDRECHT
A10
A9
A2

- **Automated process**
  - Every two weeks, data is collected and prepared, models are built and map data is generated.
  - No human work is required to do this.
  - Users can access the maps by a HTML-landing page that starts up a script that opens the desired maps.
  - Users can use CAS without needing technical knowledge.

- **CAS is leading for the planning of the following units**
  - Flexteams (entire region)
  - Teams of districts 4 and 3 (South and East)
- **For the other units, CAS is considered in the planning, but not (yet) leading.**

- **Planning process (tactical analysts):**
  - When should deployment take place?
  - Where should deployment take place?
  - Gather information on high-risk locations
    - What’s happening?
    - Who causes problems?
    - Anything else that might be interesting.
  - Production of briefing
  - Shift commences
  - Feedback after shift

- **Widening of predictive scope; also predict:**
  - Pickpocketing
  - Corporate burglary
  - Car break-in
  - Bicycle theft
- **Re-evaluate temporal component**
- **Migrate to different mapping tool**

Questions?