Anomaly Detection and Training Data Scarcity

W. Robertson
UC Berkeley

F. Maggi
Politecnico di Milano

C. Kruegel, G. Vigna
UC Santa Barbara

NDSS 2010
San Diego, CA

Anomaly detection

- Black-box, “hands-free” approach for detecting attacks
- Profiles constructed from training set
- Deviations from profiles considered to be attacks

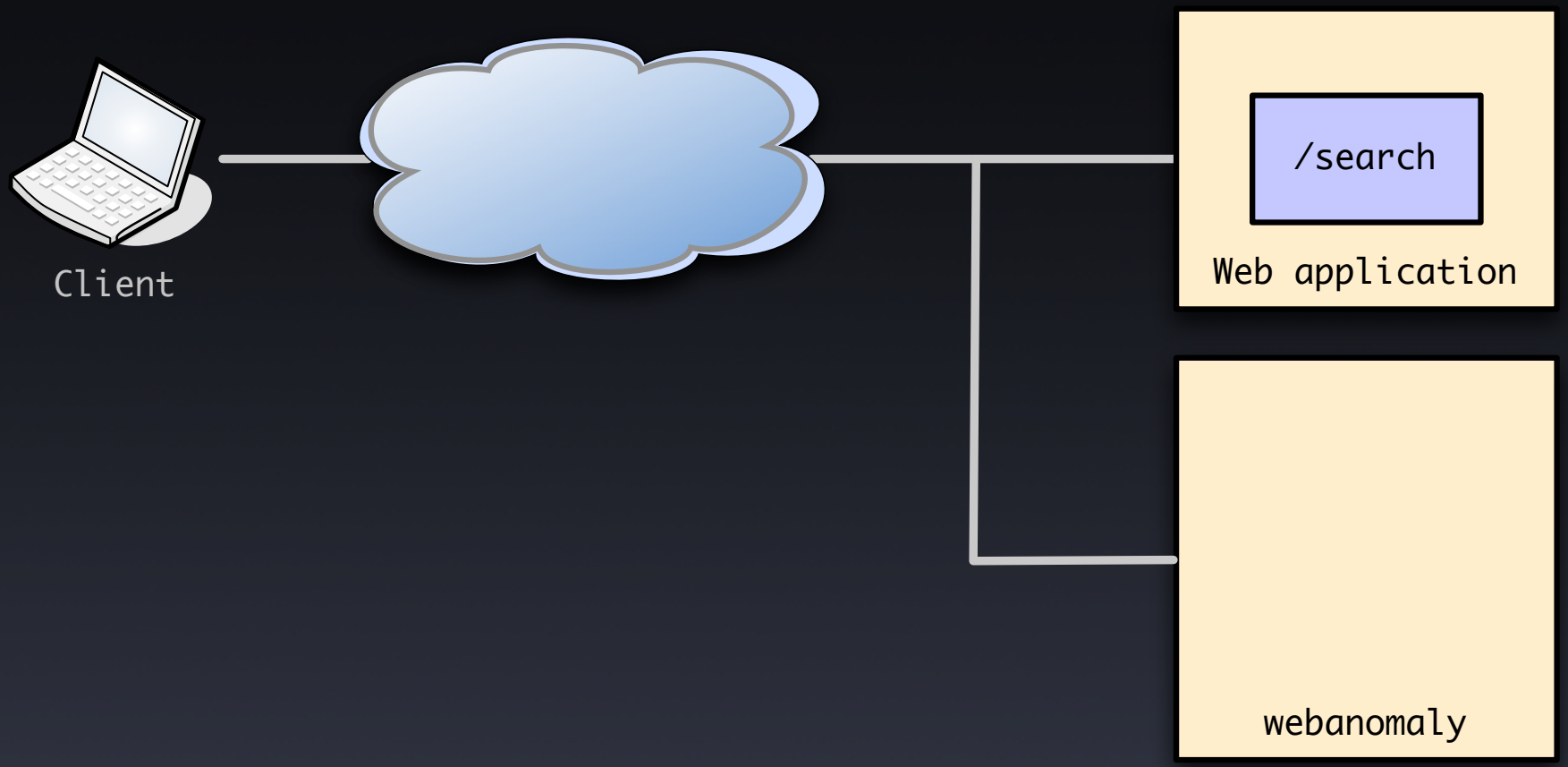
Anomaly detection

- Black-box, “hands-free” approach for detecting attacks
- Profiles constructed from training set
- Deviations from profiles considered to be attacks
- Focus on web application anomaly detection

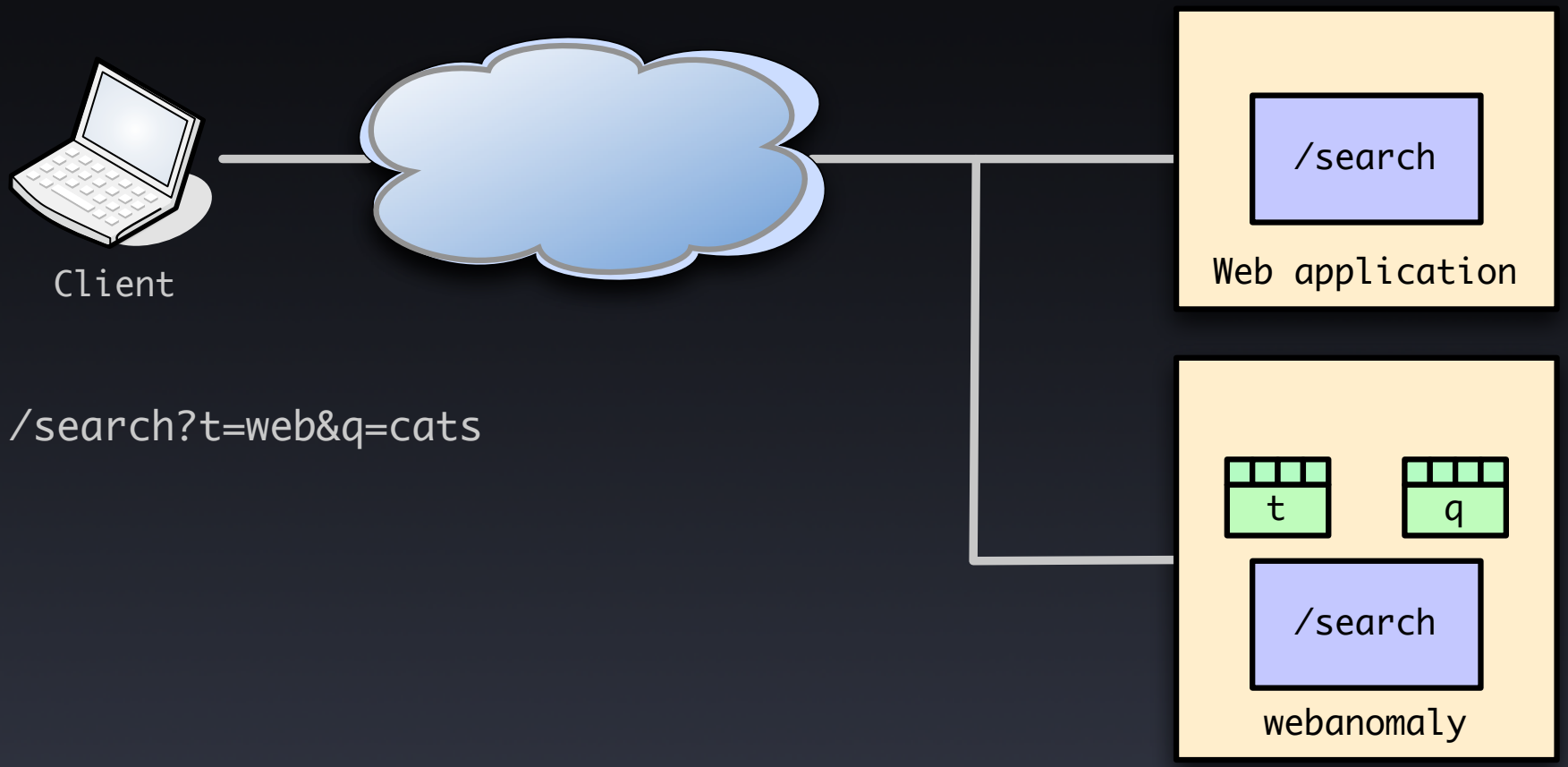
Training data

- Detection quality crucially depends on training data
- Data can be noisy
- Data can be incomplete

Ex: Web application

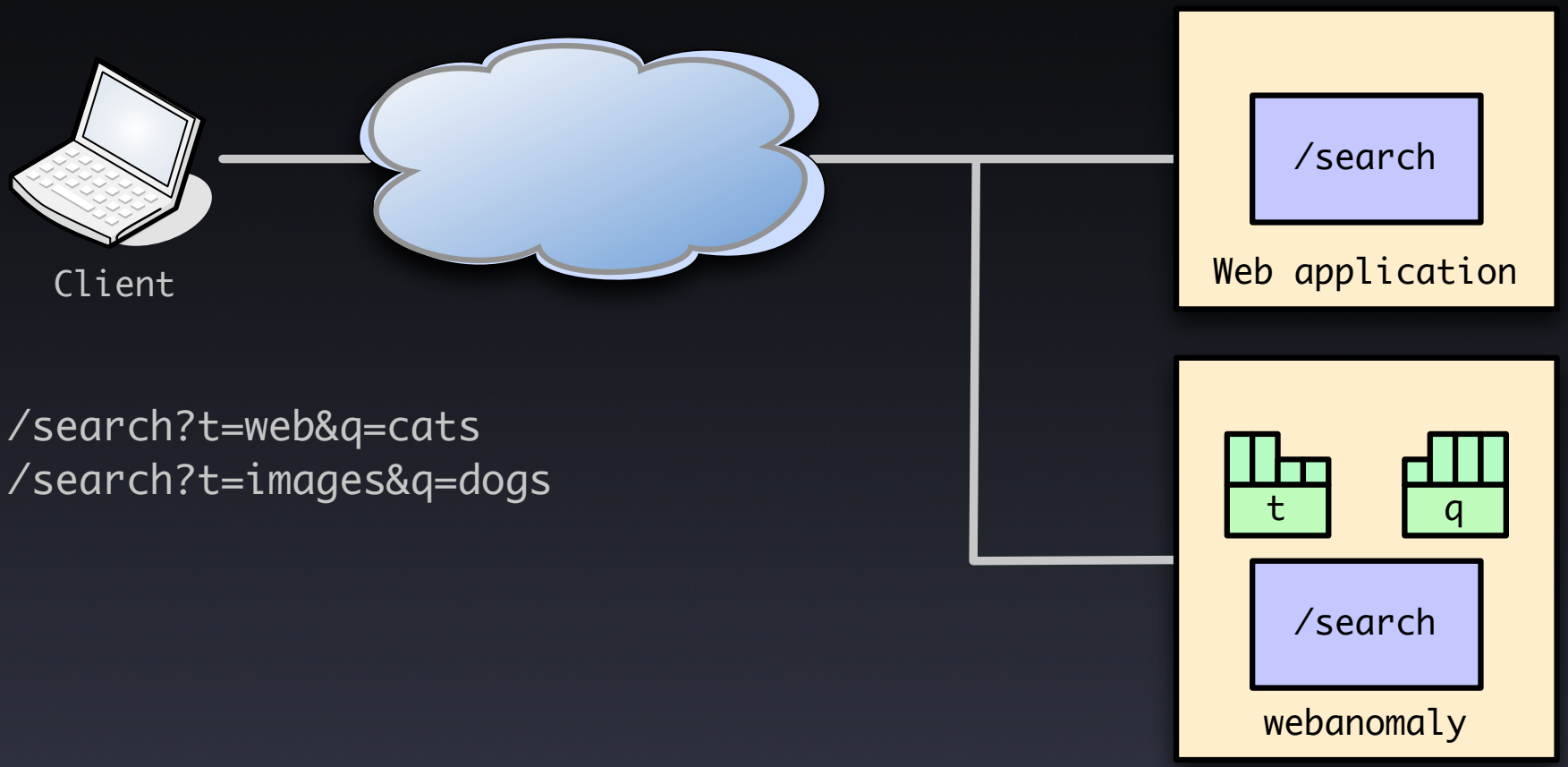


Ex: Web application



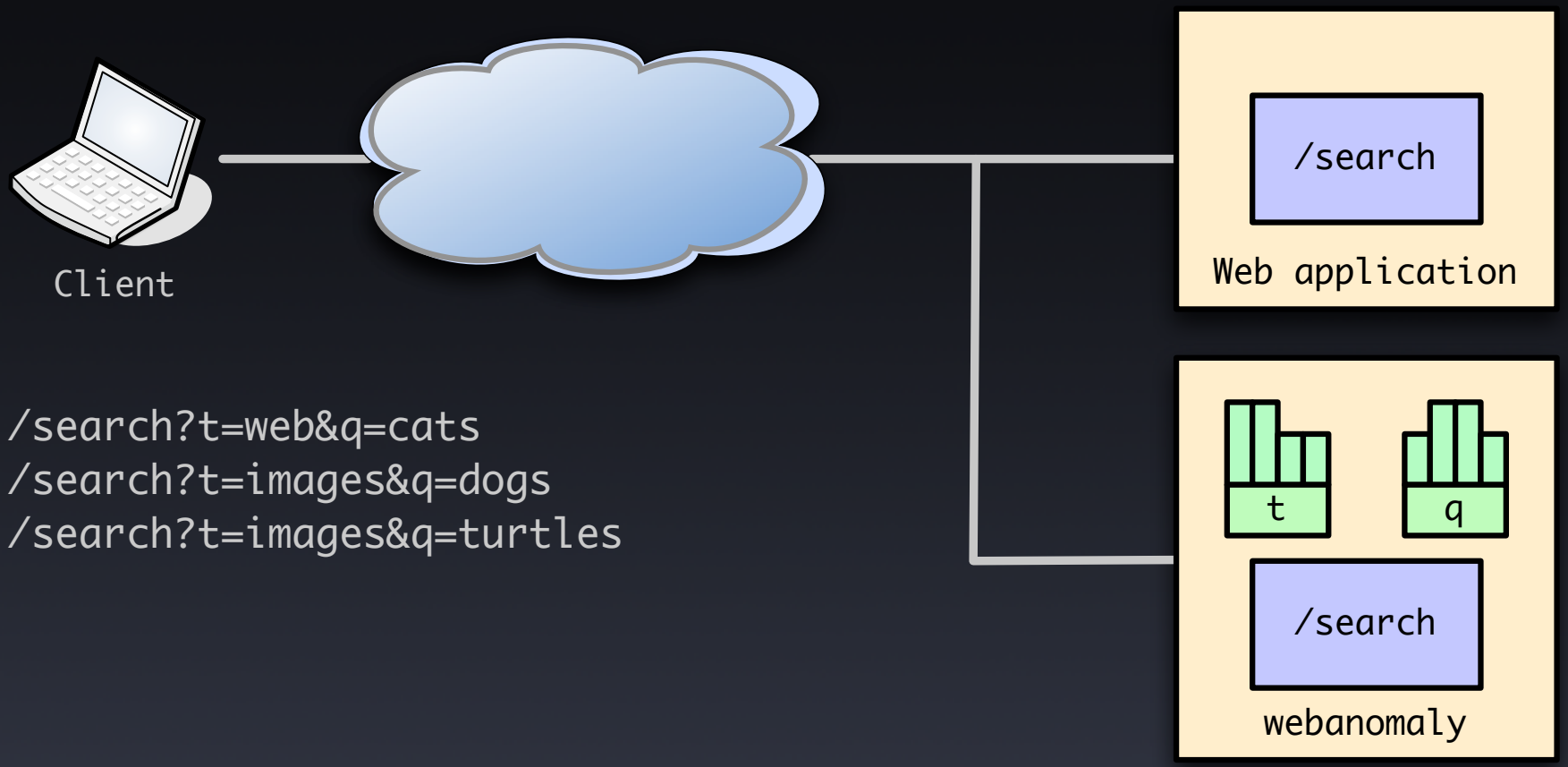
/search?t=web&q=cats

Ex: Web application



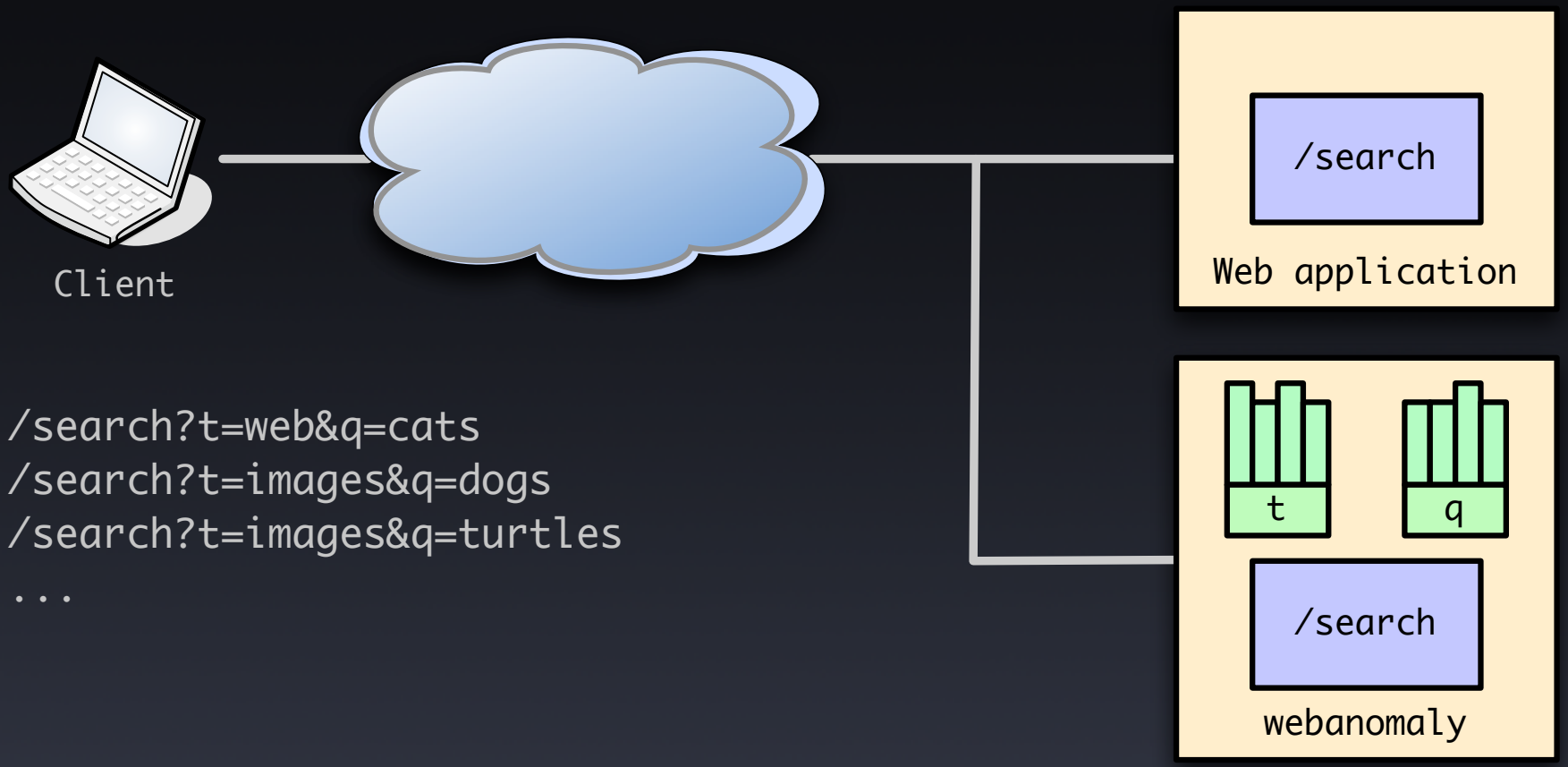

```
/search?t=web&q=cats
/search?t=images&q=dogs
```

Ex: Web application



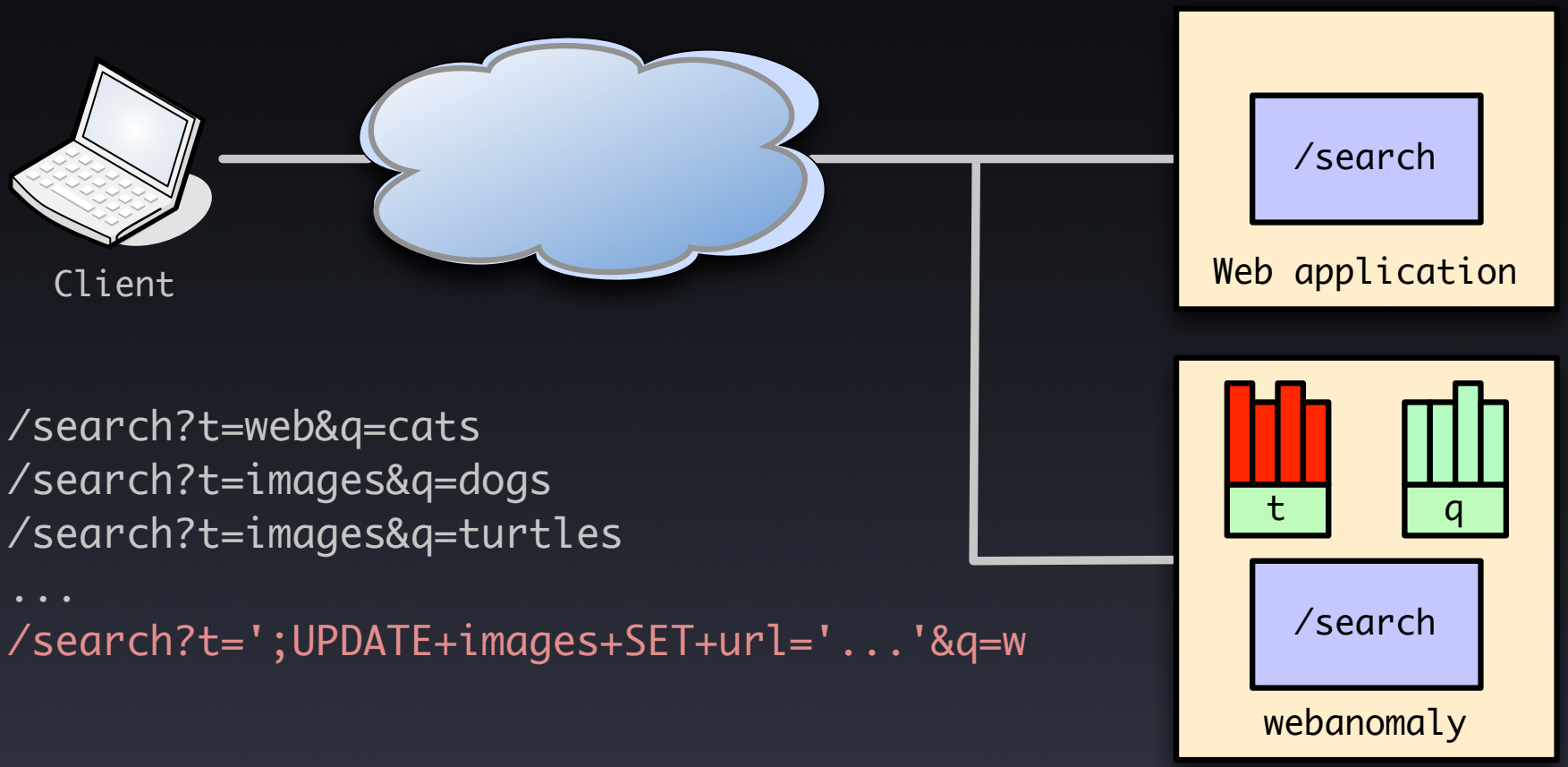

```
/search?t=web&q=cats
/search?t=images&q=dogs
/search?t=images&q=turtles
```

Ex: Web application



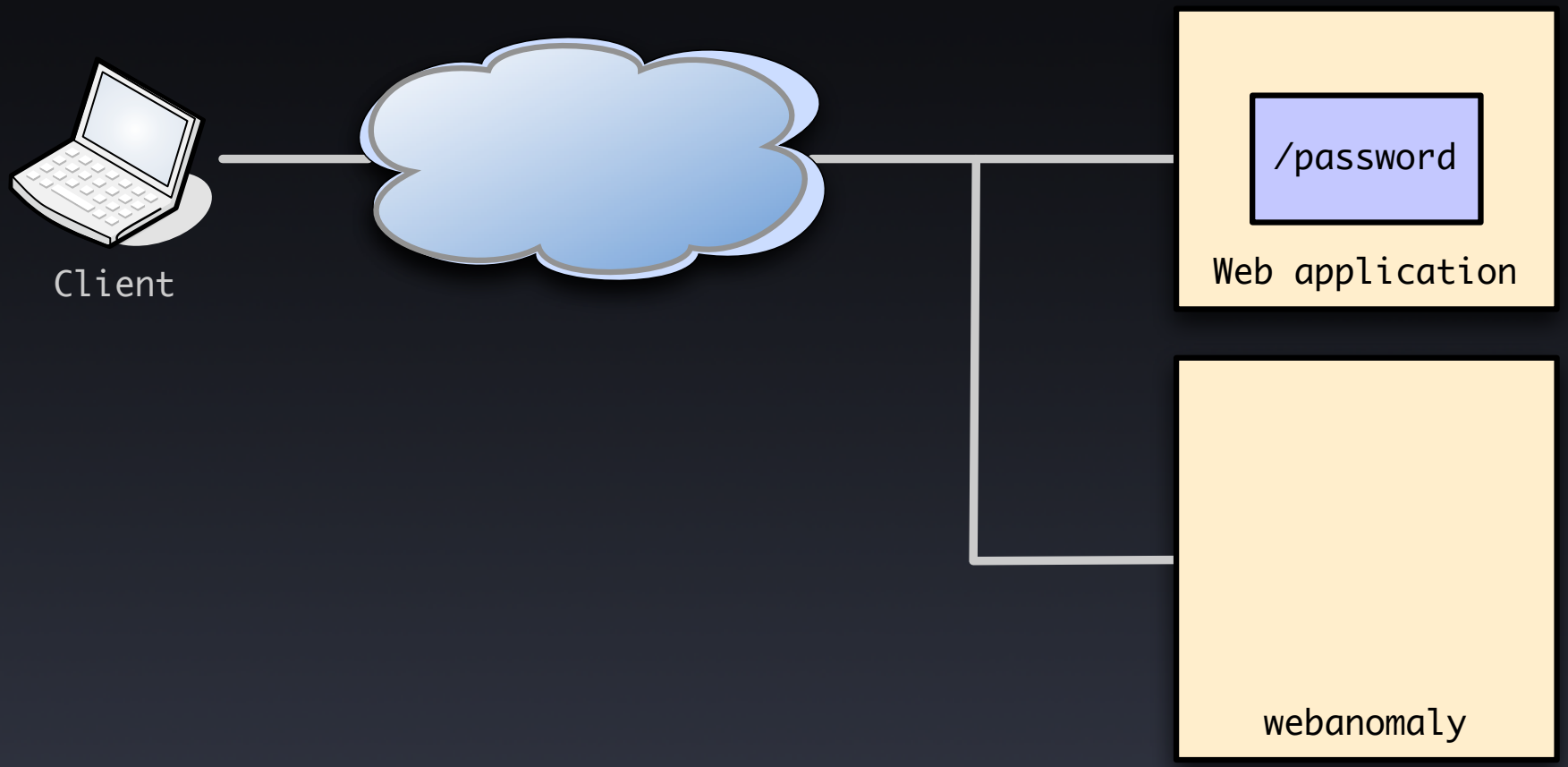

```
/search?t=web&q=cats
/search?t=images&q=dogs
/search?t=images&q=turtles
...
```

Ex: Web application

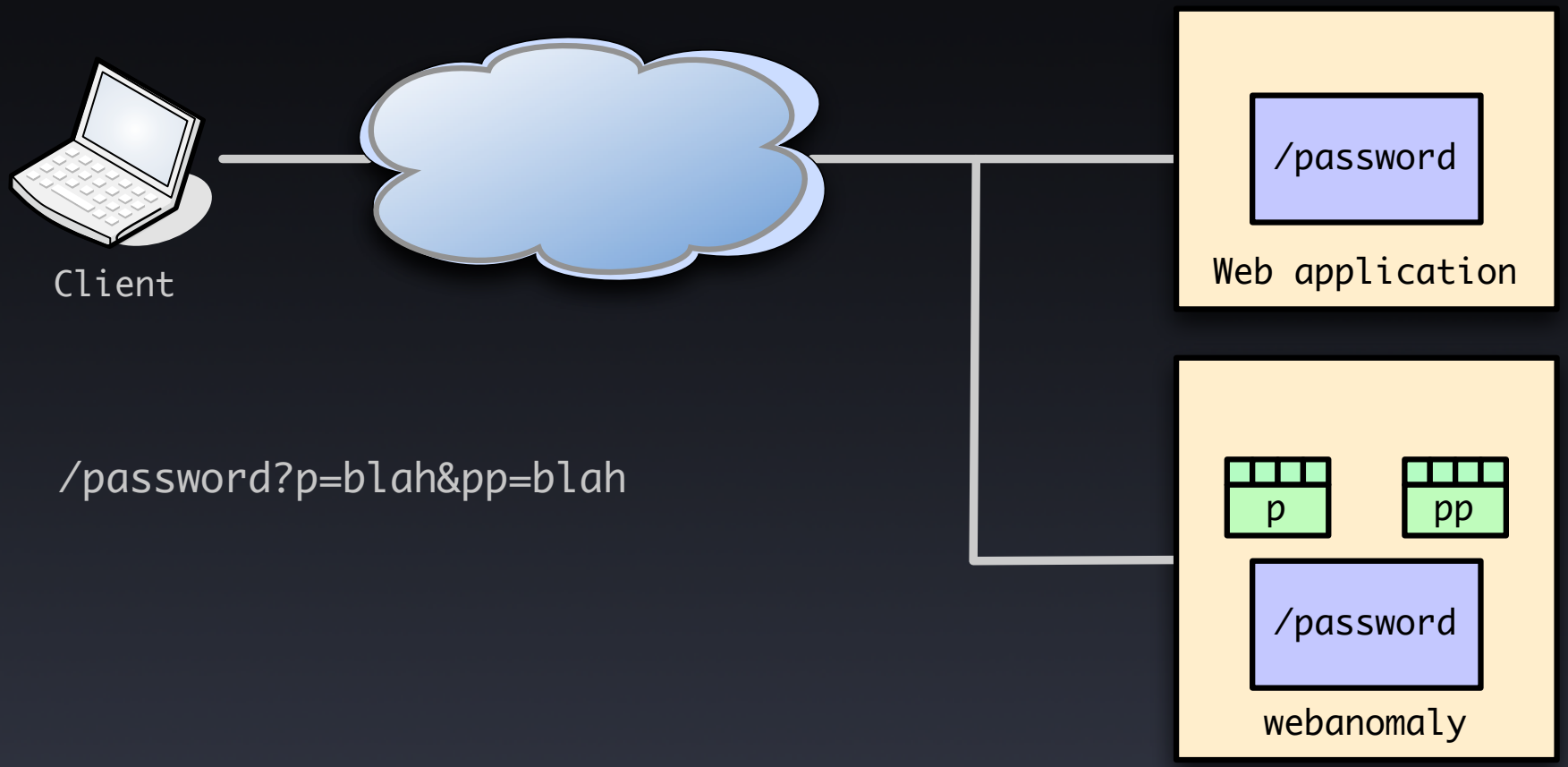



```
/search?t=web&q=cats
/search?t=images&q=dogs
/search?t=images&q=turtles
...
/search?t=';UPDATE+images+SET+url='...'&q=w
```

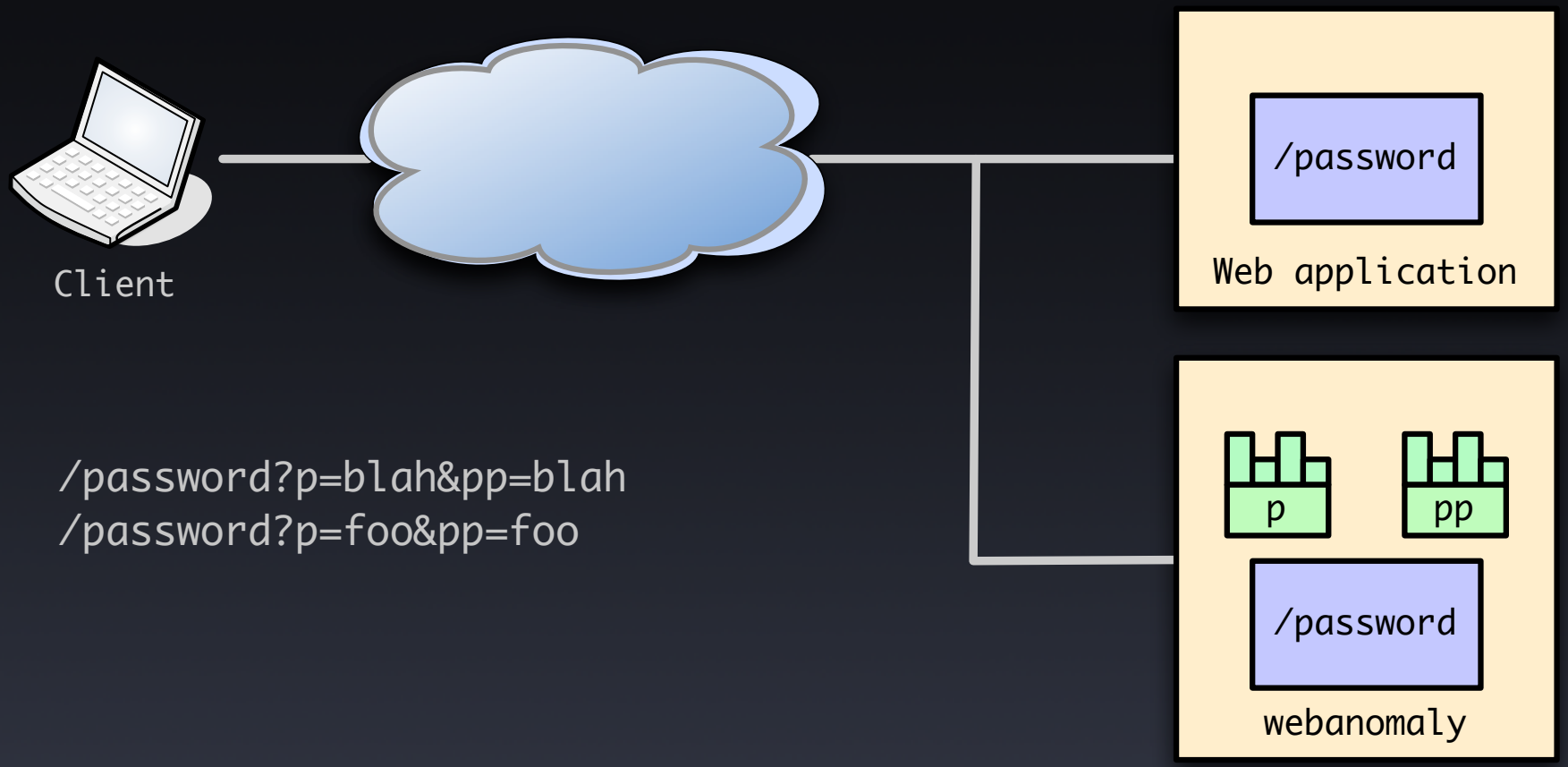
Ex: Web application



Ex: Web application

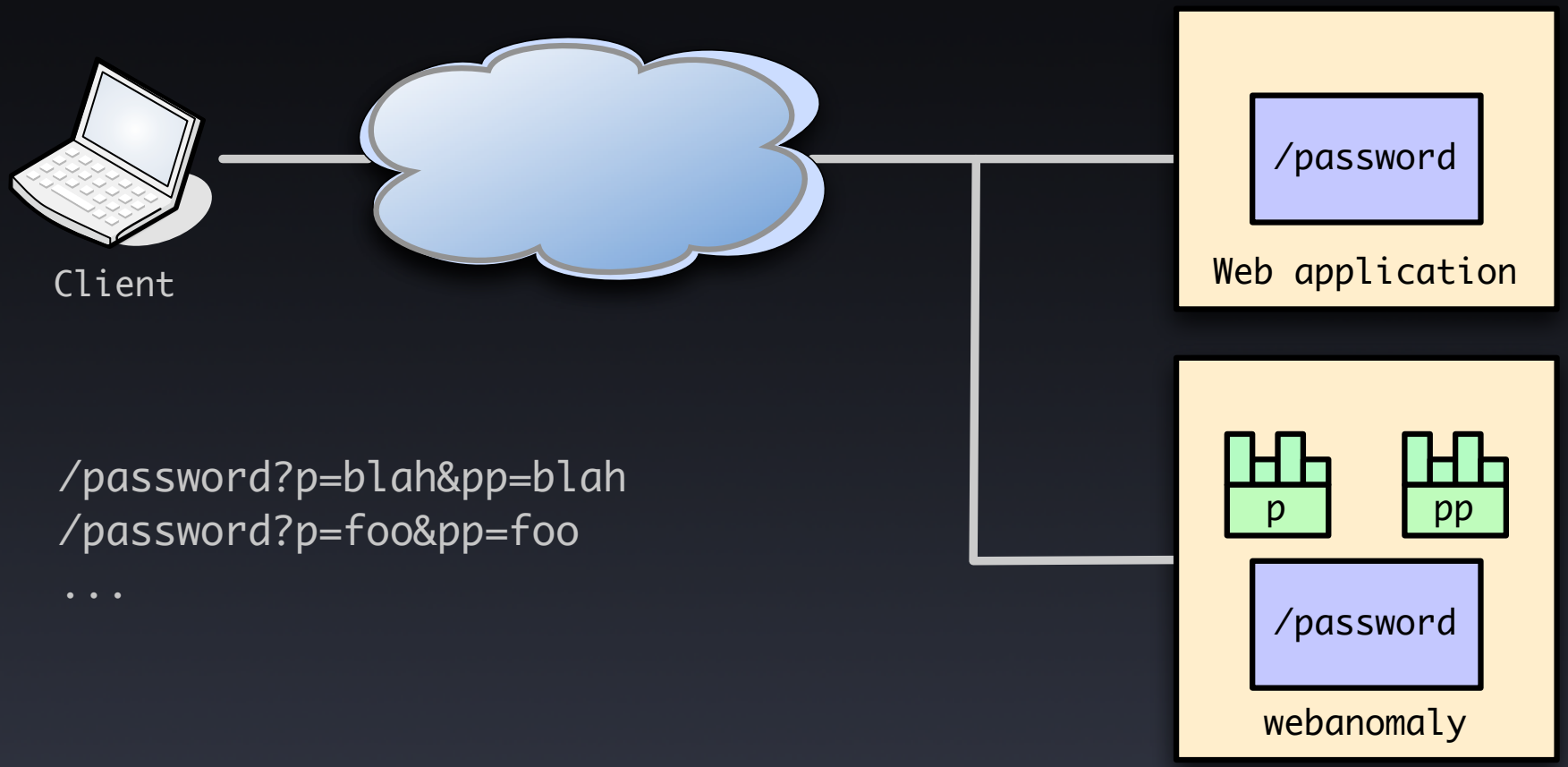


Ex: Web application



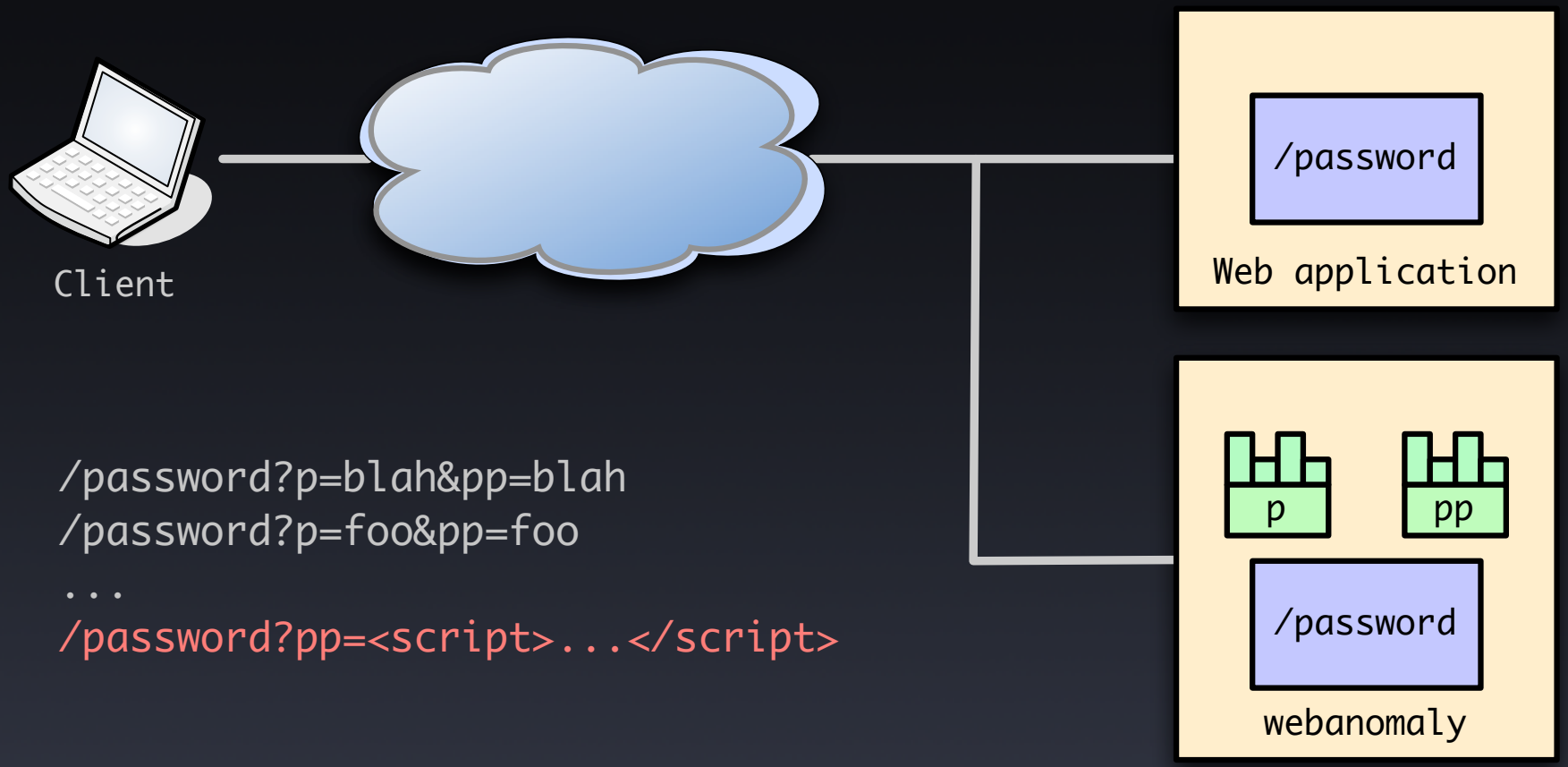

```
/password?p=blah&pp=blah
/password?p=foo&pp=foo
```

Ex: Web application



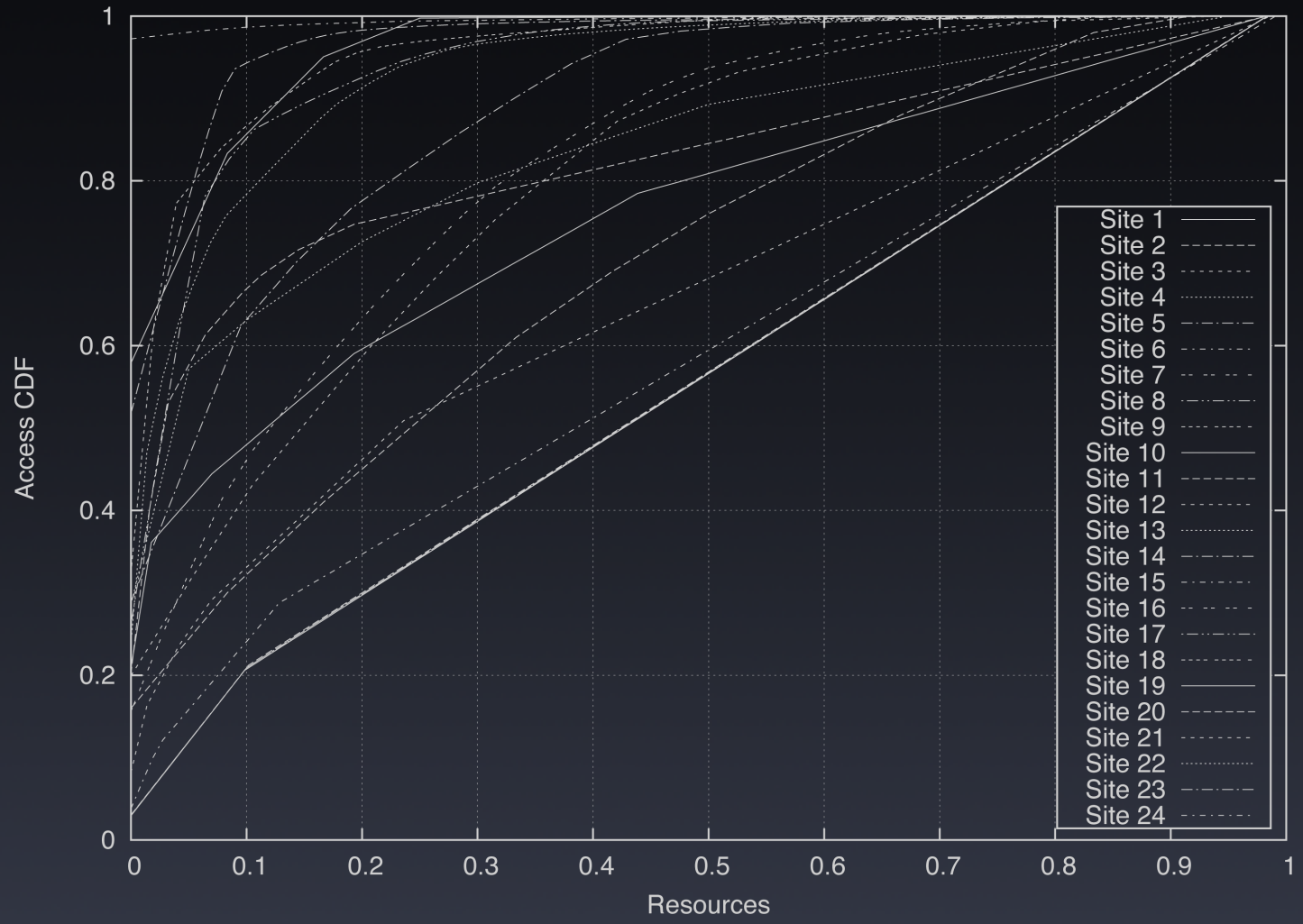

```
/password?p=blah&pp=blah
/password?p=foo&pp=foo
...
```

Ex: Web application




```
/password?p=blah&pp=blah
/password?p=foo&pp=foo
...
/password?pp=<script>...</script>
```

Resource invocation distribution



Overview

Introduction

Exploiting global information

- Enhanced training phase
- Under-trained profile database

Using global profiles

Evaluation

Conclusions

Profile clusters

Profile clusters

Profile clusters

Under-trained profiles

Observation

1. Features often can be grouped into semantically-similar clusters.

2. Similar features induce similar profiles.

Under-trained profiles

Observation

1. Features often can be grouped into semantically-similar clusters.

2. Similar features induce similar profiles.

Can under-trained profiles be replaced by similar well-trained profiles?

Under-trained profiles

cats
dogs
turtles
...

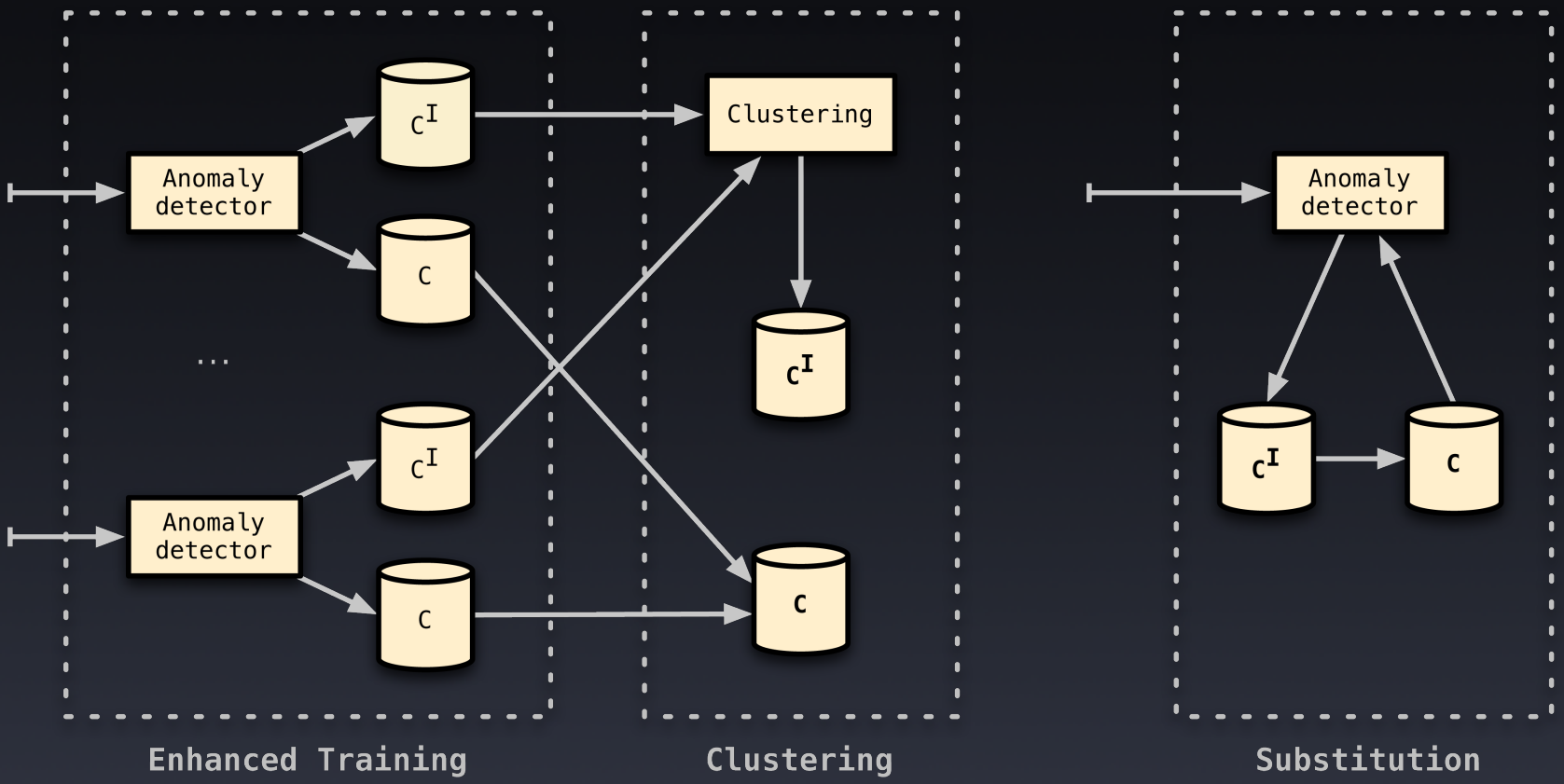
blah
foo

Under-trained profiles

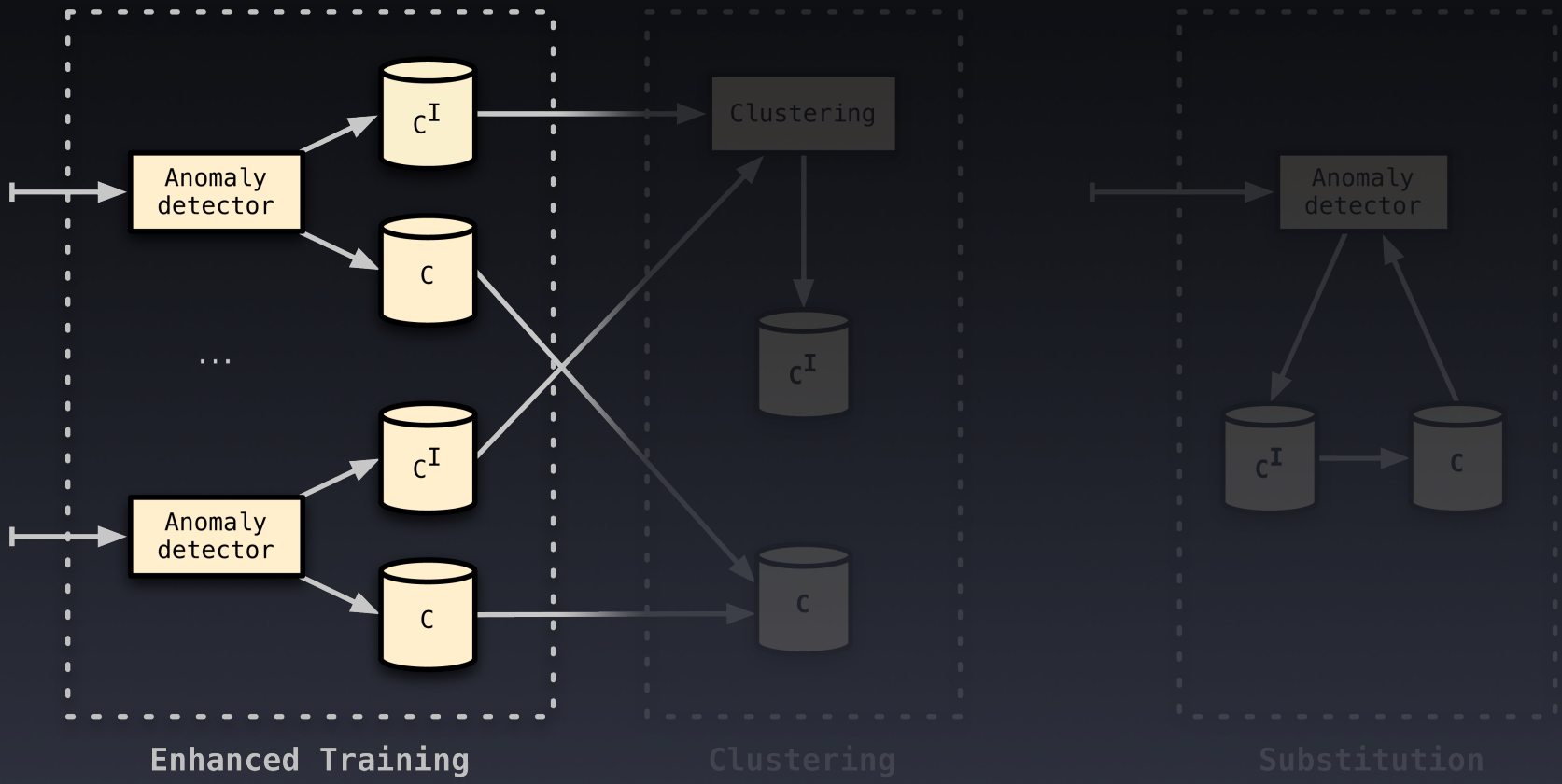
cats
dogs

blah
foo

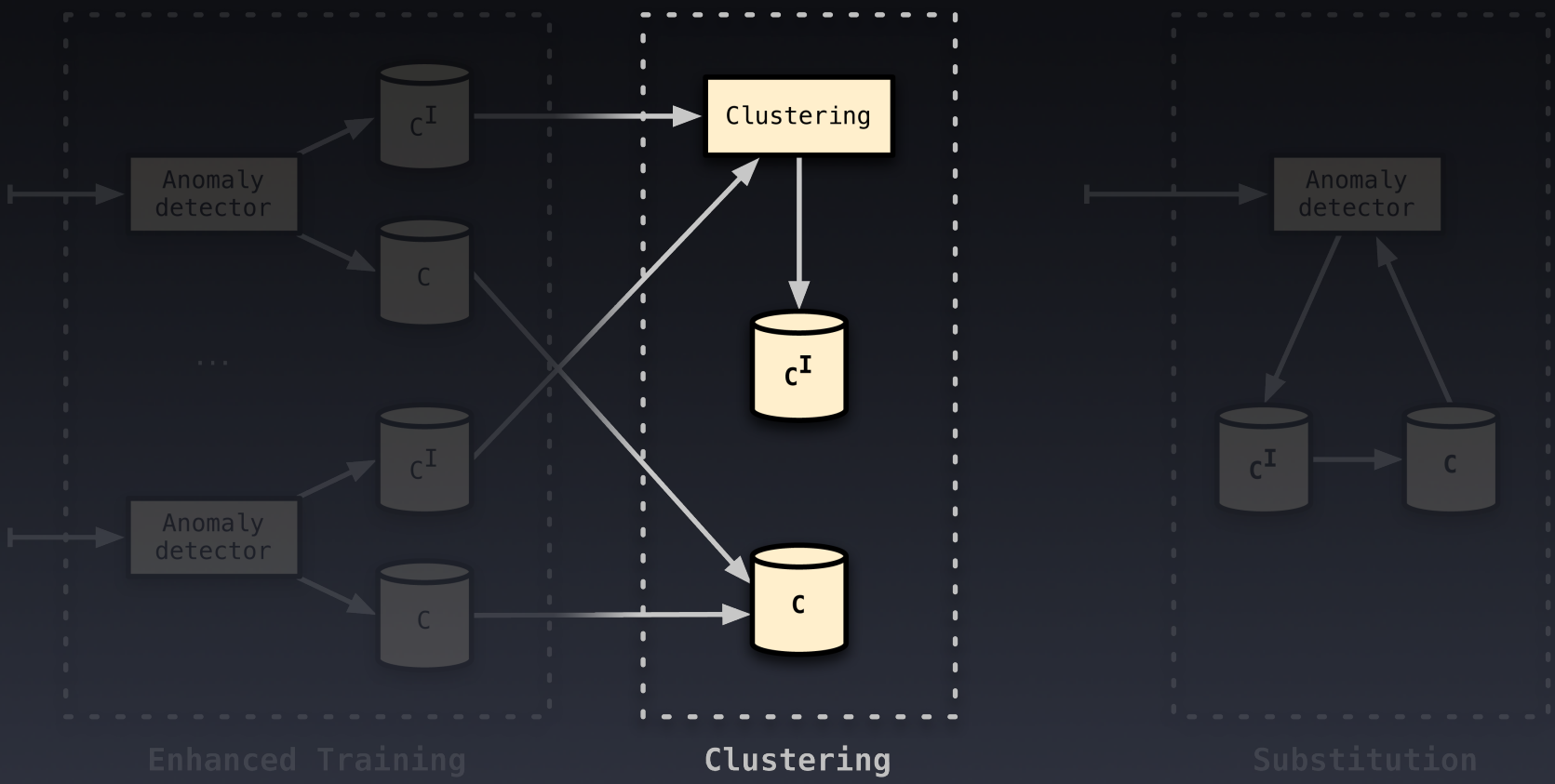
Overall procedure



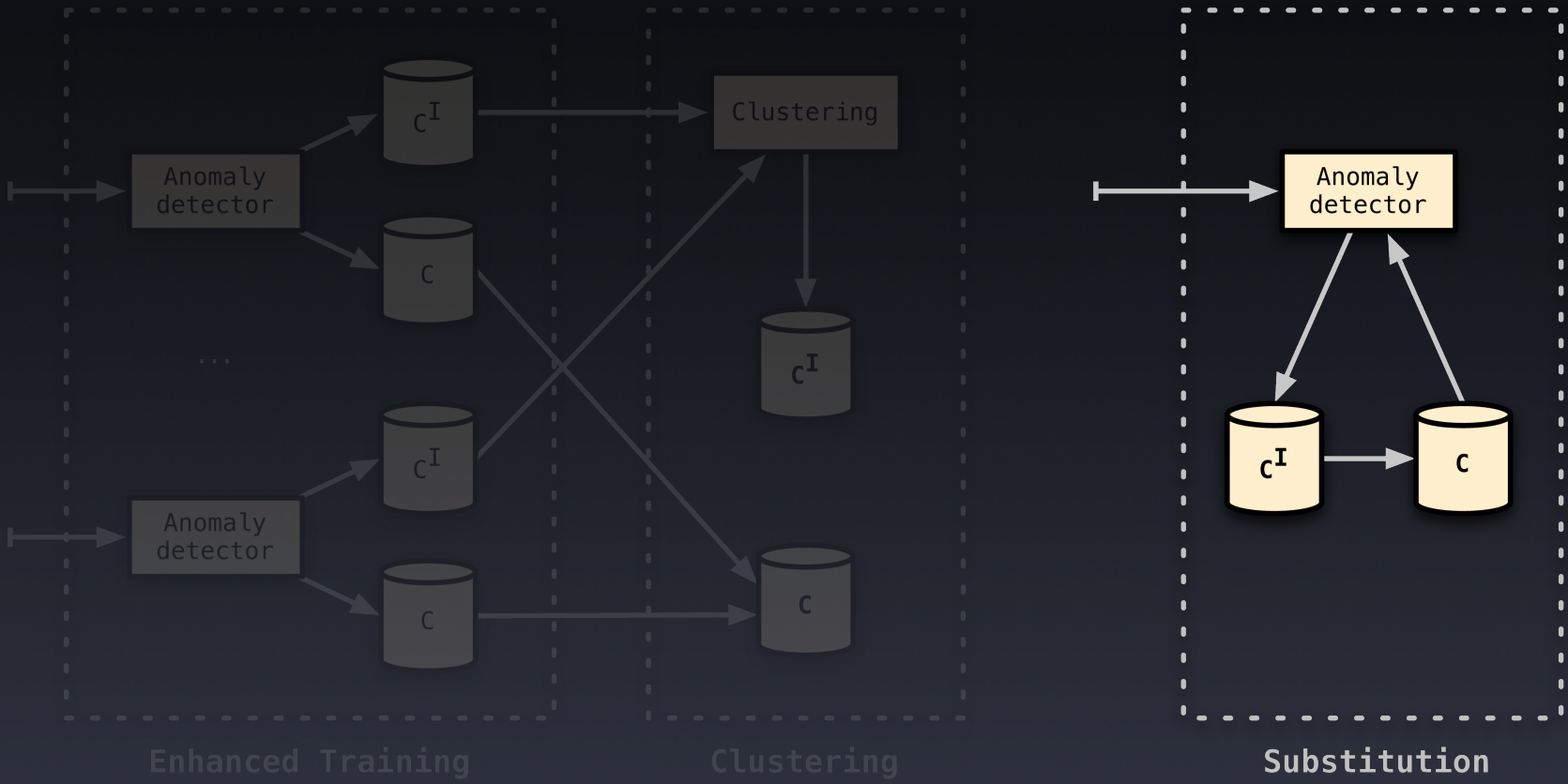
Overall procedure



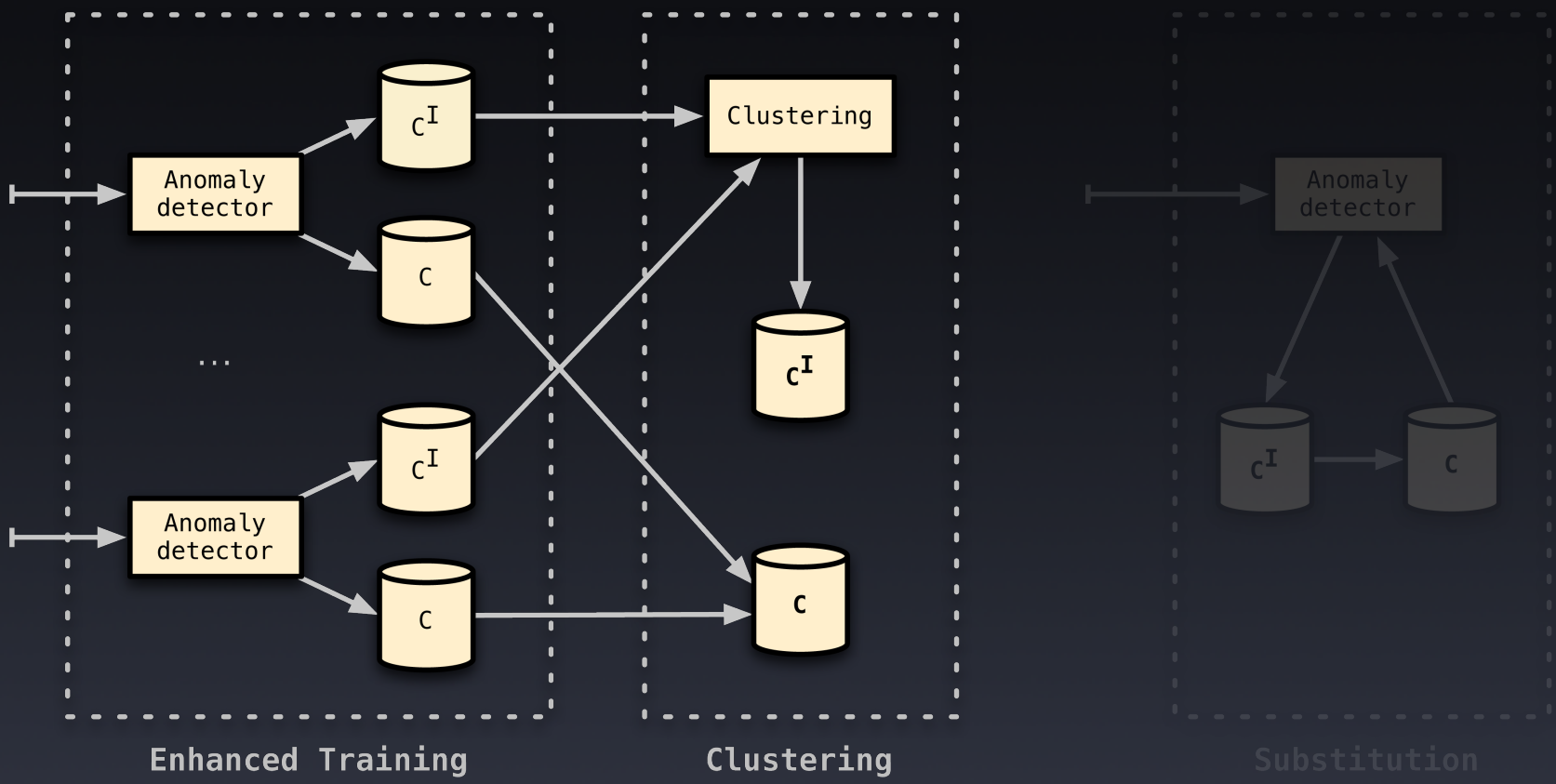
Overall procedure



Overall procedure

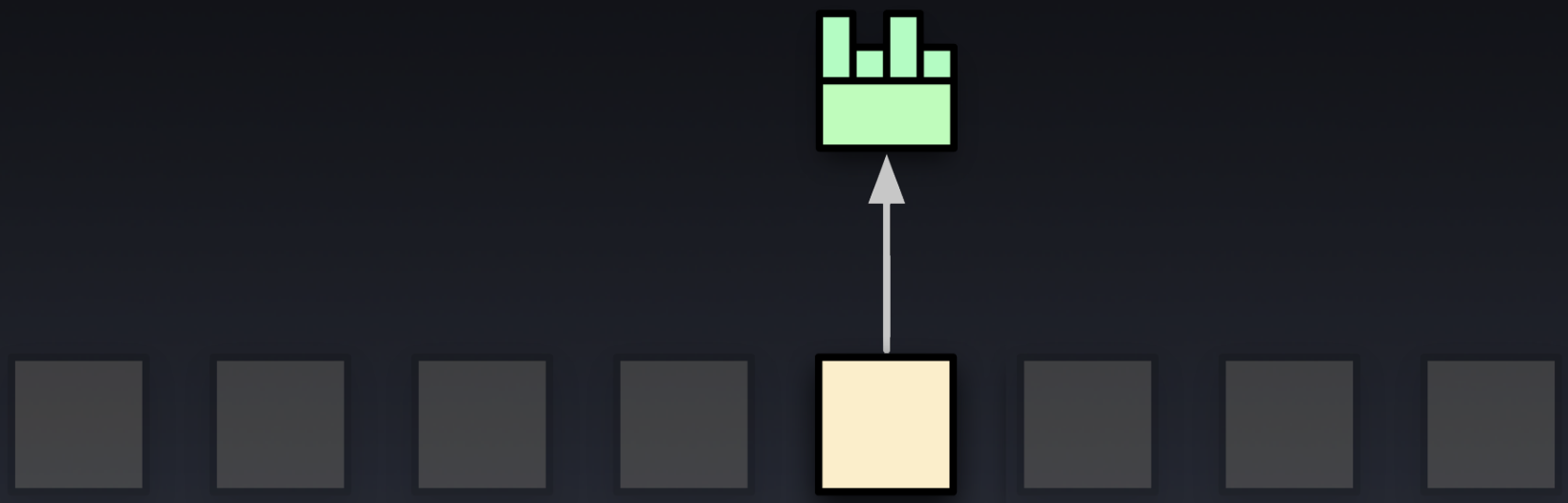


Training procedure



Sampling under-trained profiles

Sampling under-trained profiles

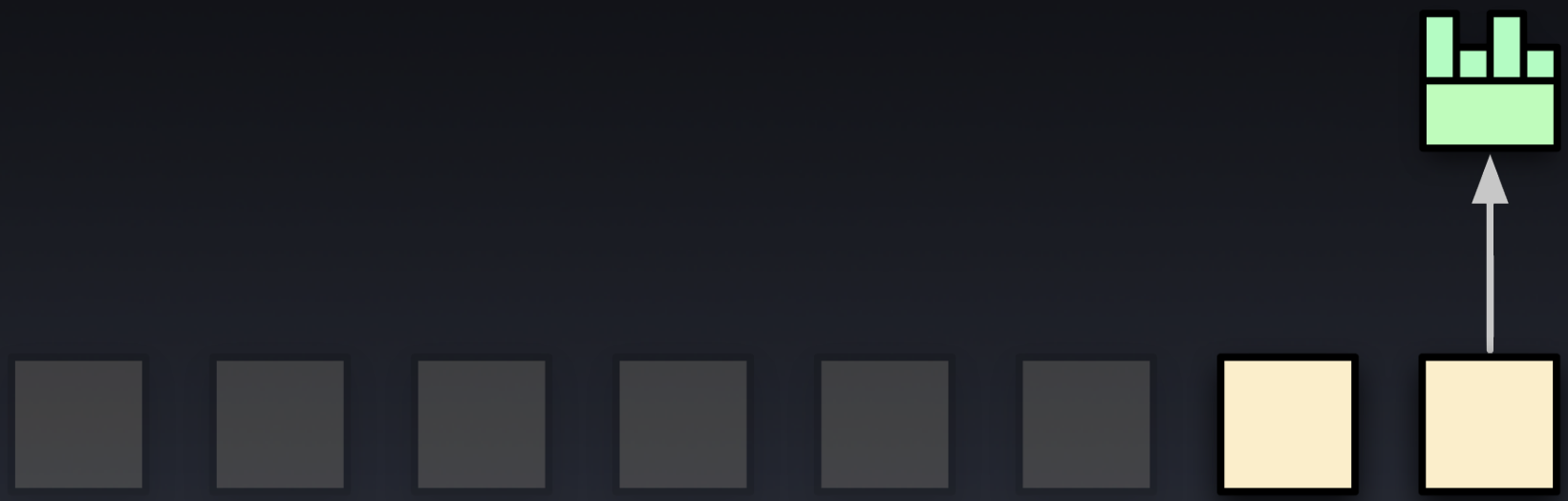


$$\kappa = 1$$

Sampling under-trained profiles

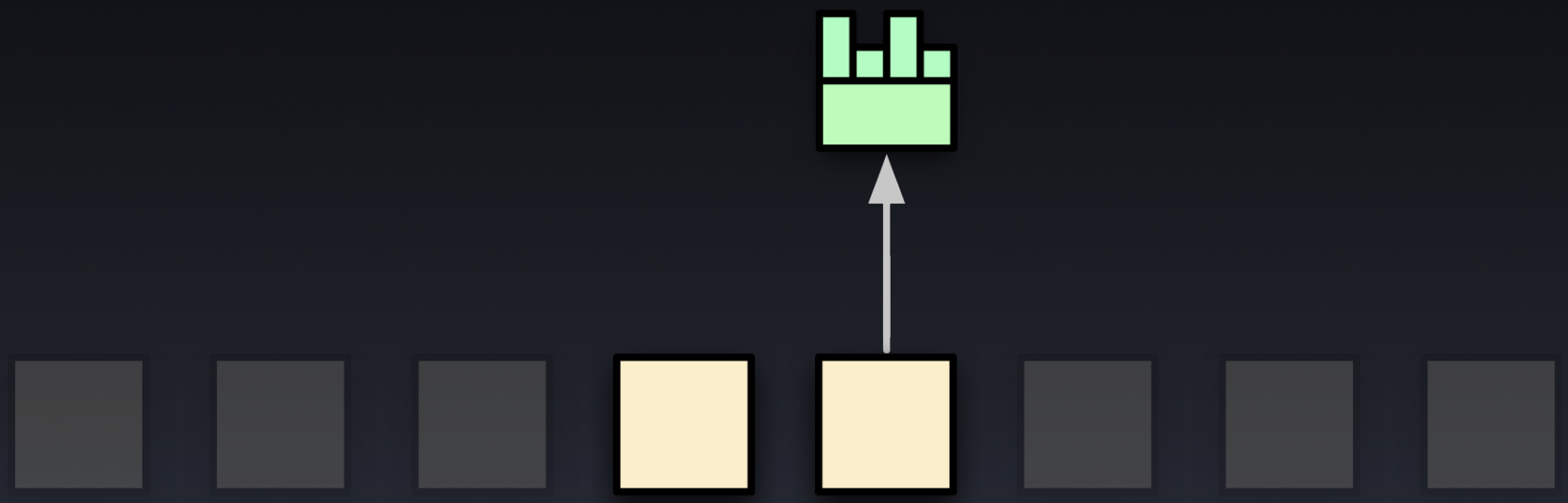
$$\kappa = 1$$

Sampling under-trained profiles



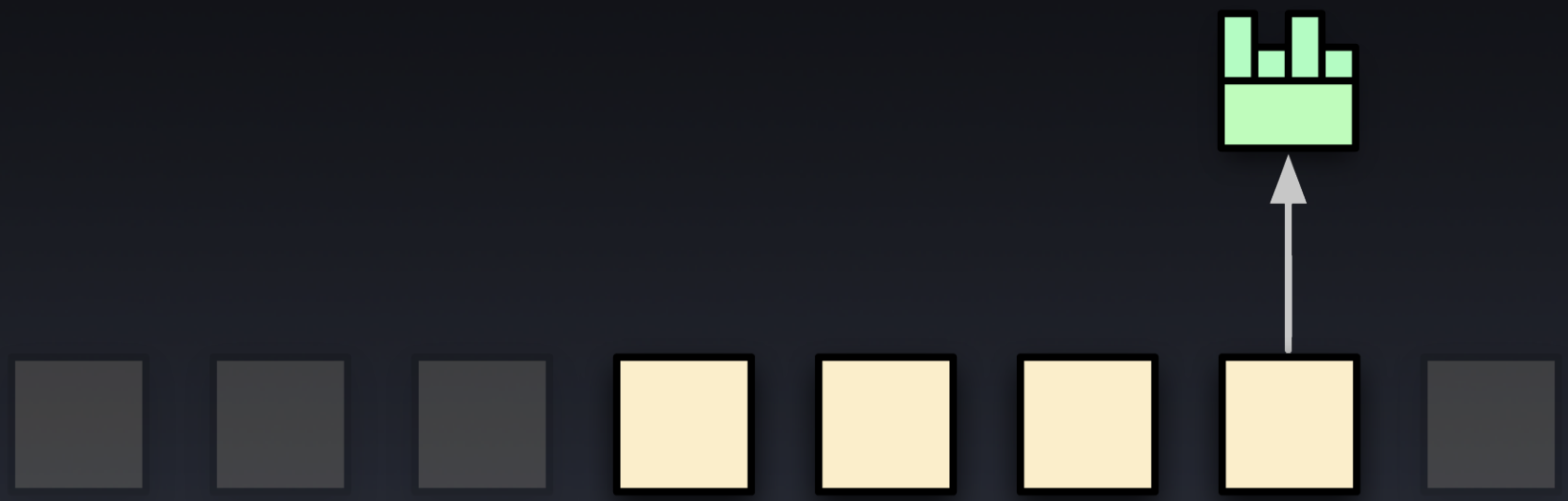
$$\kappa = 2$$

Sampling under-trained profiles



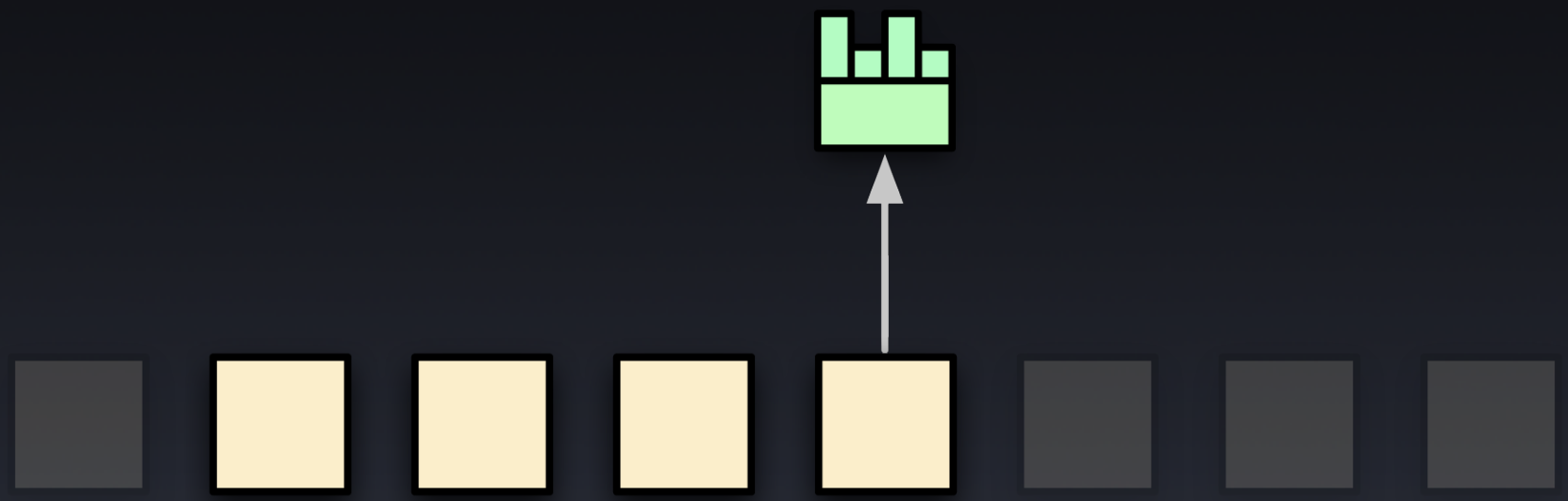
$$\kappa = 2$$

Sampling under-trained profiles



$$\kappa = 4$$

Sampling under-trained profiles



$$\kappa = 4$$

Clustering under-trained profiles

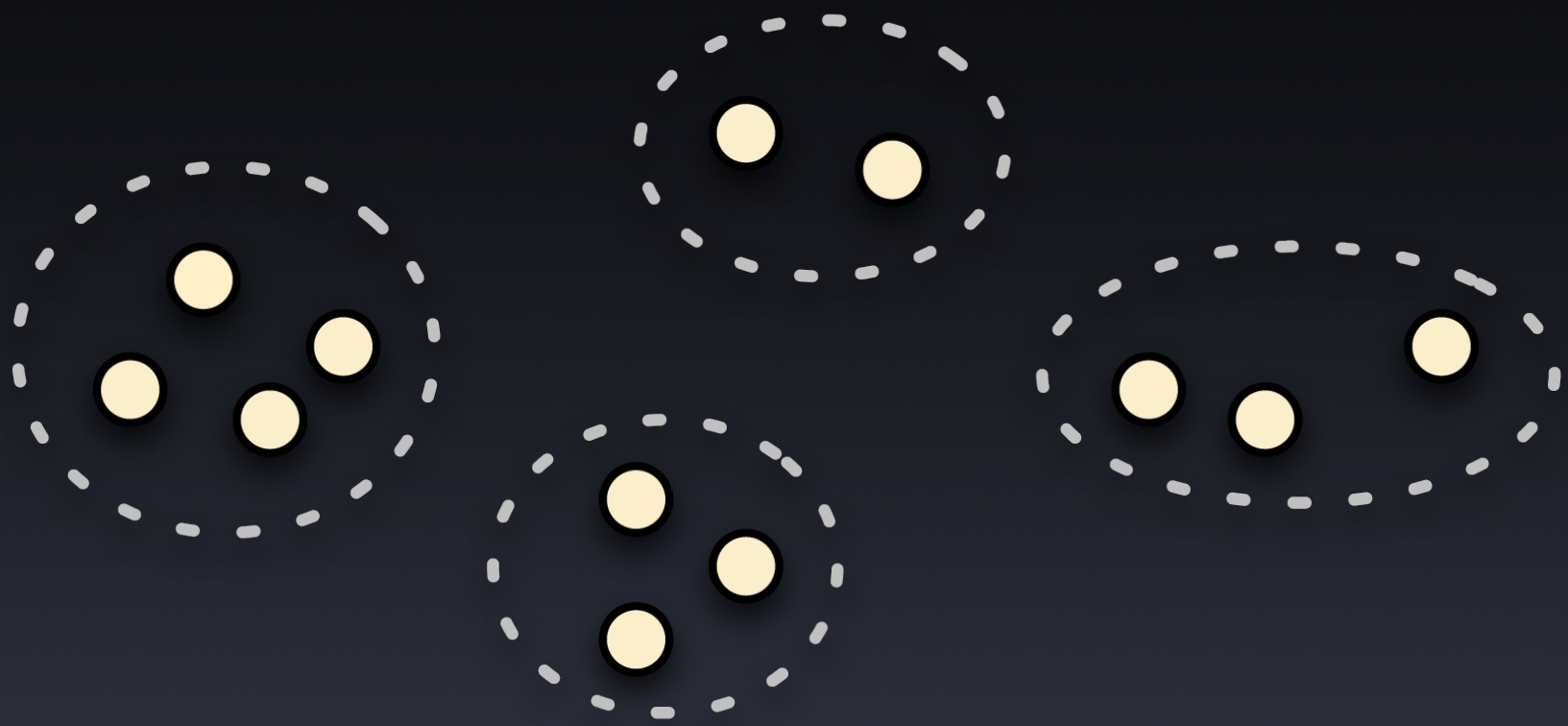
- Profile distance function

$$\delta\left(c_i, c_j\right) = \frac{1}{|c_i \cap c_j|} \sum_{m_i^{(\cdot)}, m_j^{(\cdot)} \in c_i \cap c_j} \delta_{(\cdot)}\left(m_i^{(\cdot)}, m_j^{(\cdot)}\right)$$

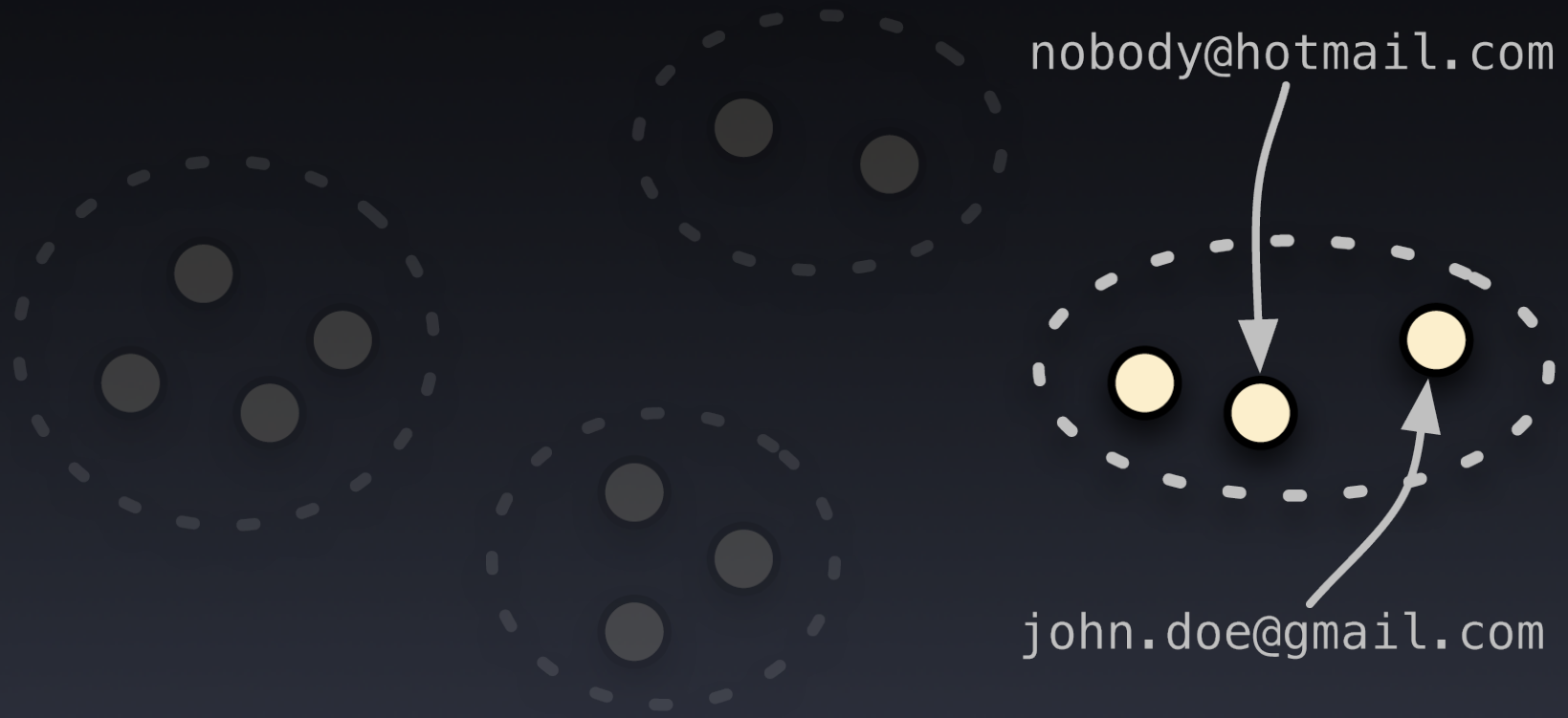
- Structure model distance function

$$\delta_s\left(m_i^s, m_j^s\right) = 1 - \frac{\mathbb{O}_i \cap \mathbb{O}_j}{\mathbb{O}_i \cup \mathbb{O}_j}$$

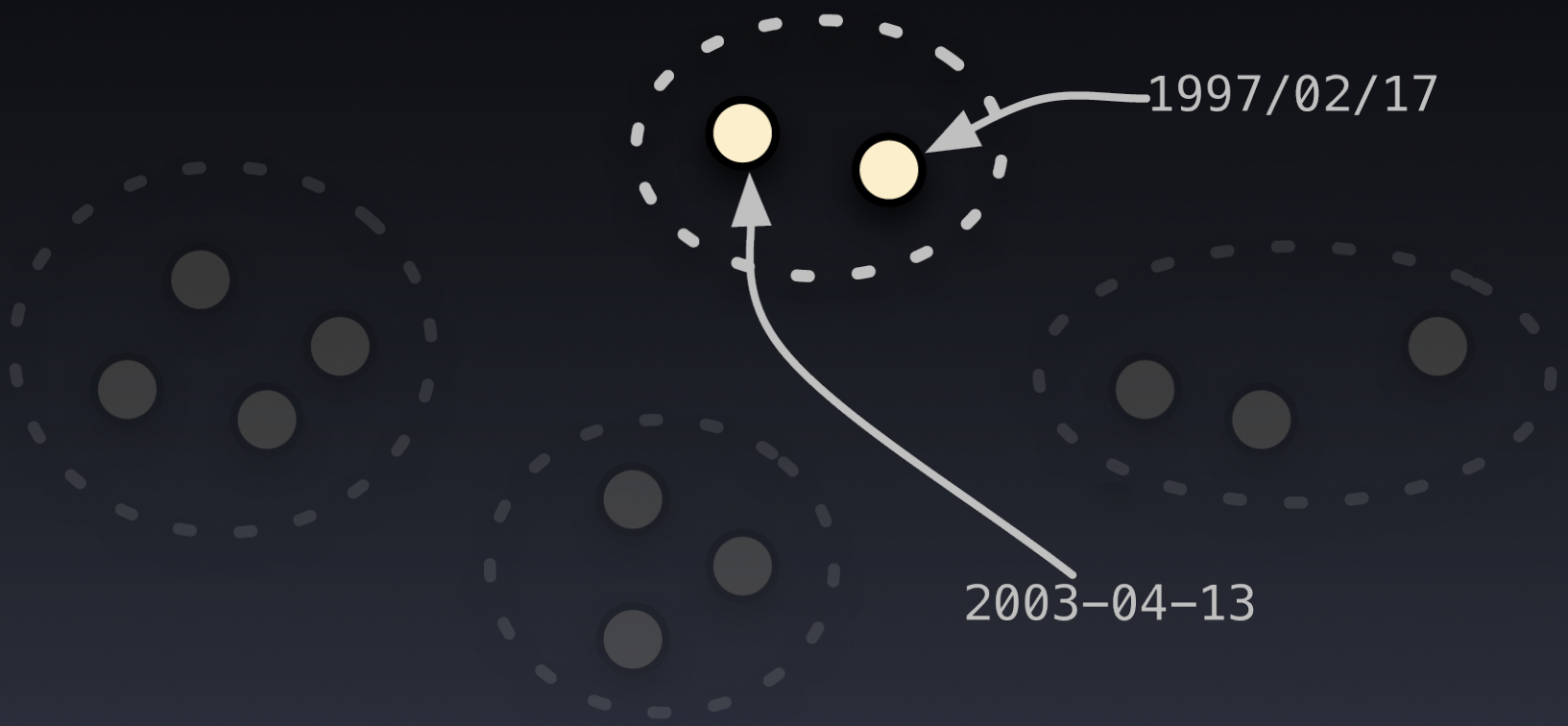
Clustering under-trained profiles



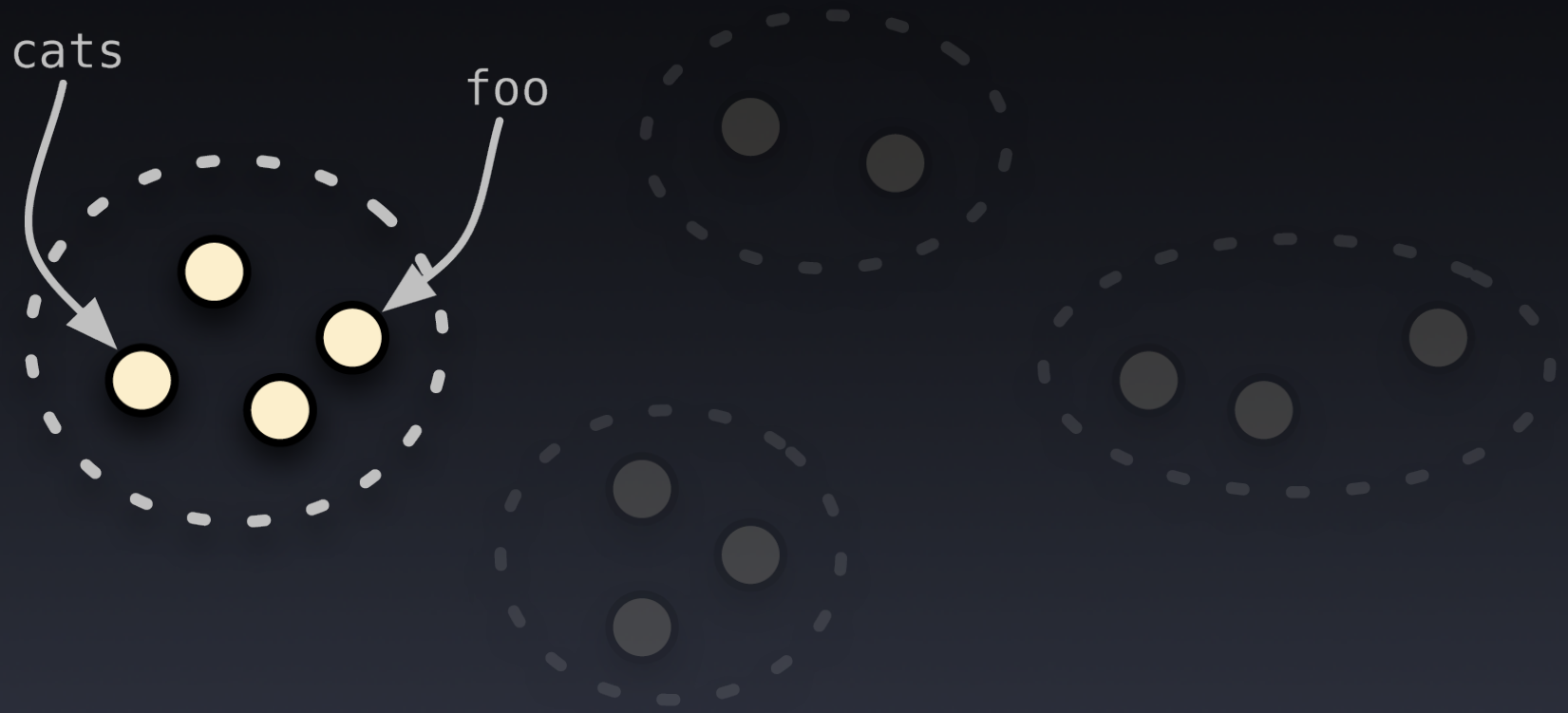
Clustering under-trained profiles



Clustering under-trained profiles

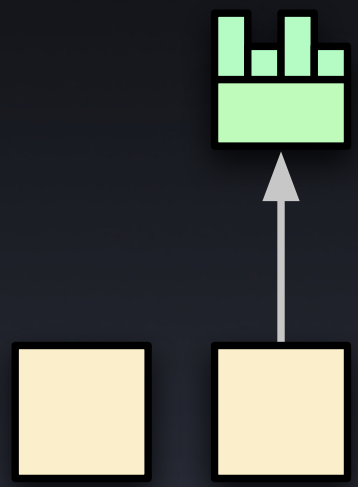


Clustering under-trained profiles

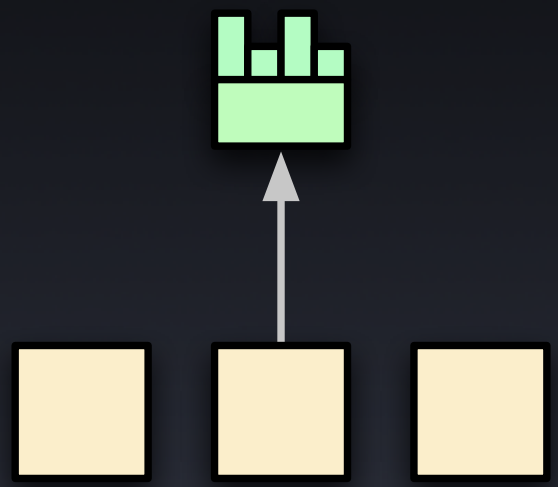


Mapping under-trained profiles

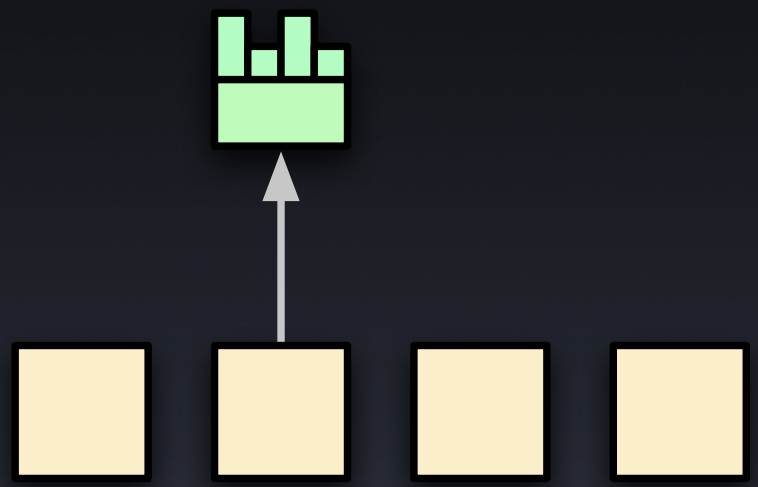
Mapping under-trained profiles



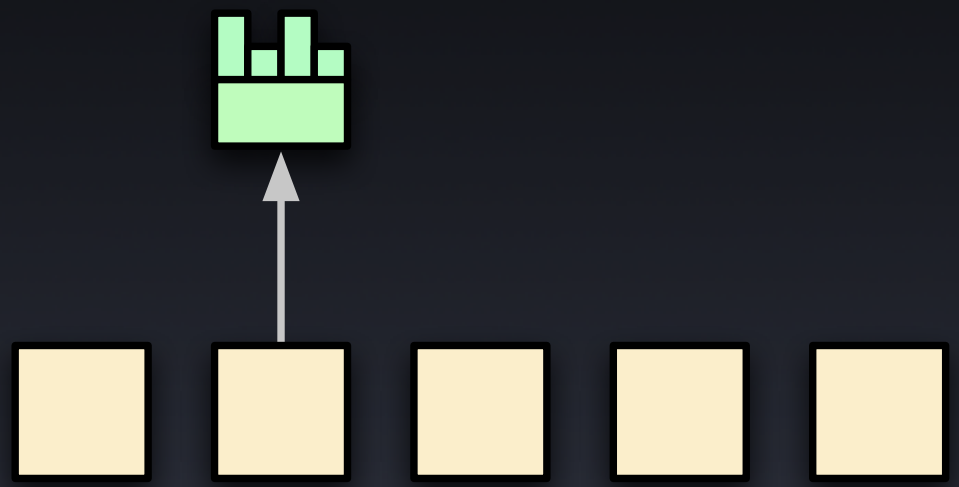
Mapping under-trained profiles



Mapping under-trained profiles



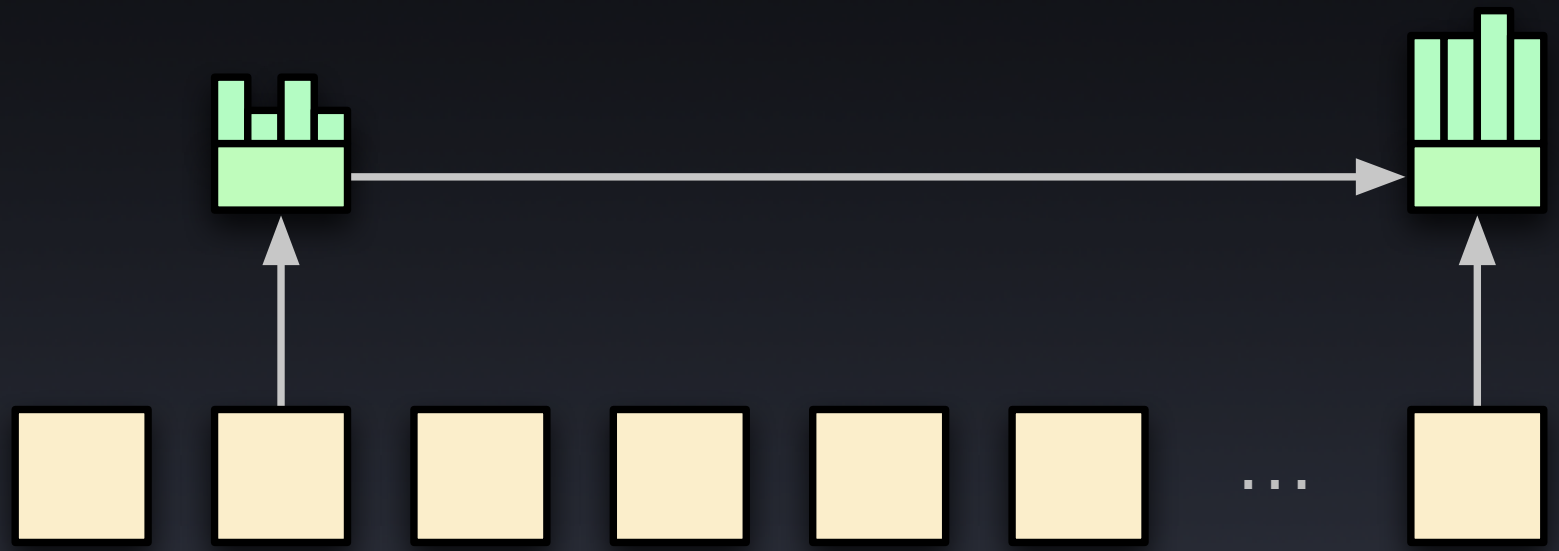
Mapping under-trained profiles



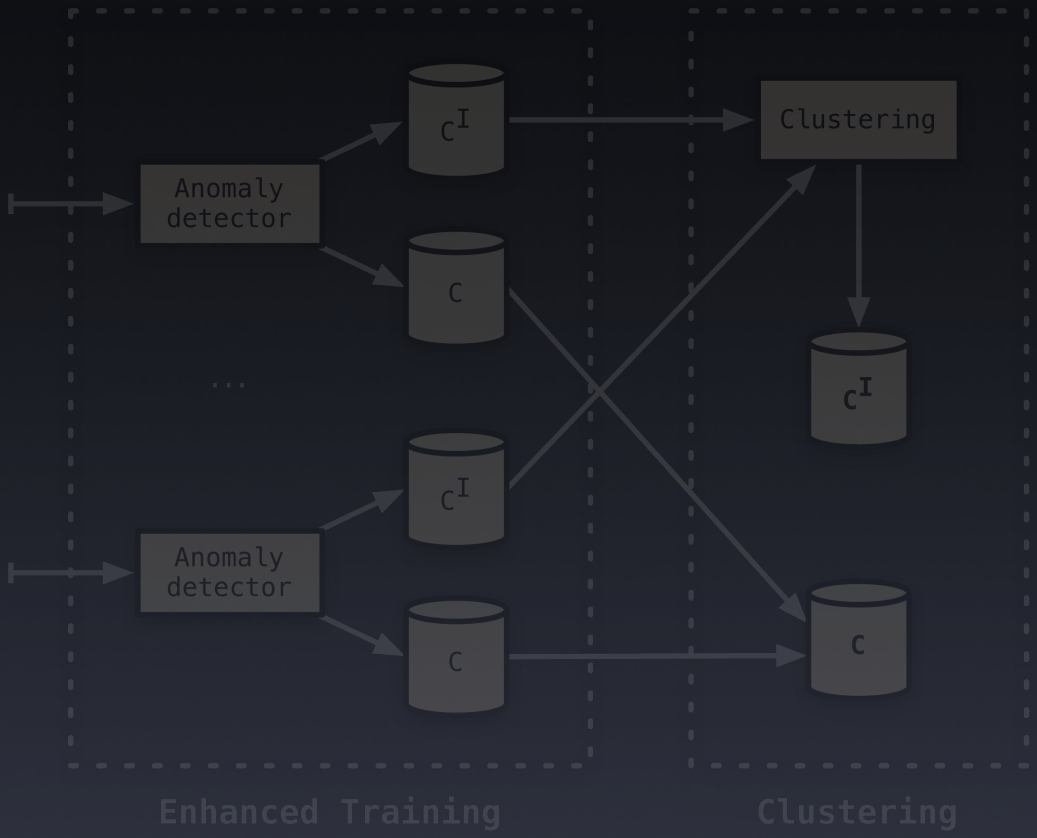
Mapping under-trained profiles

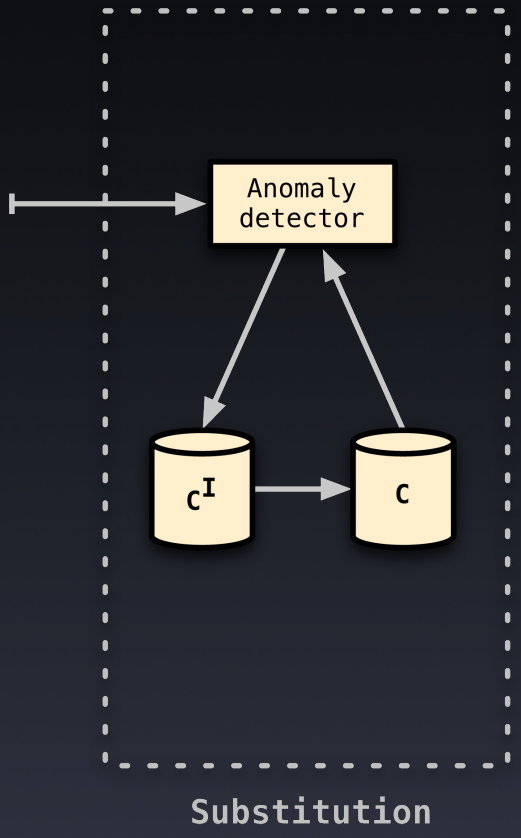
Mapping under-trained profiles

Mapping under-trained profiles

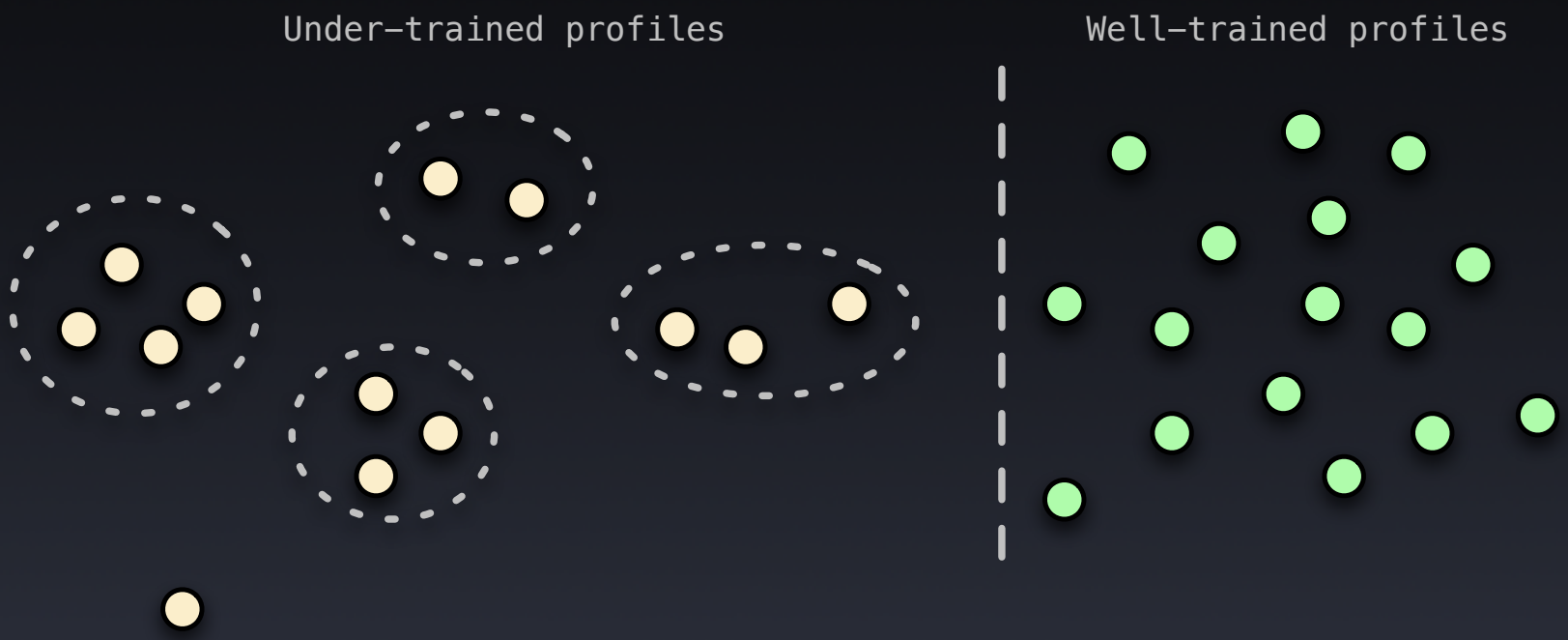


Substitution procedure

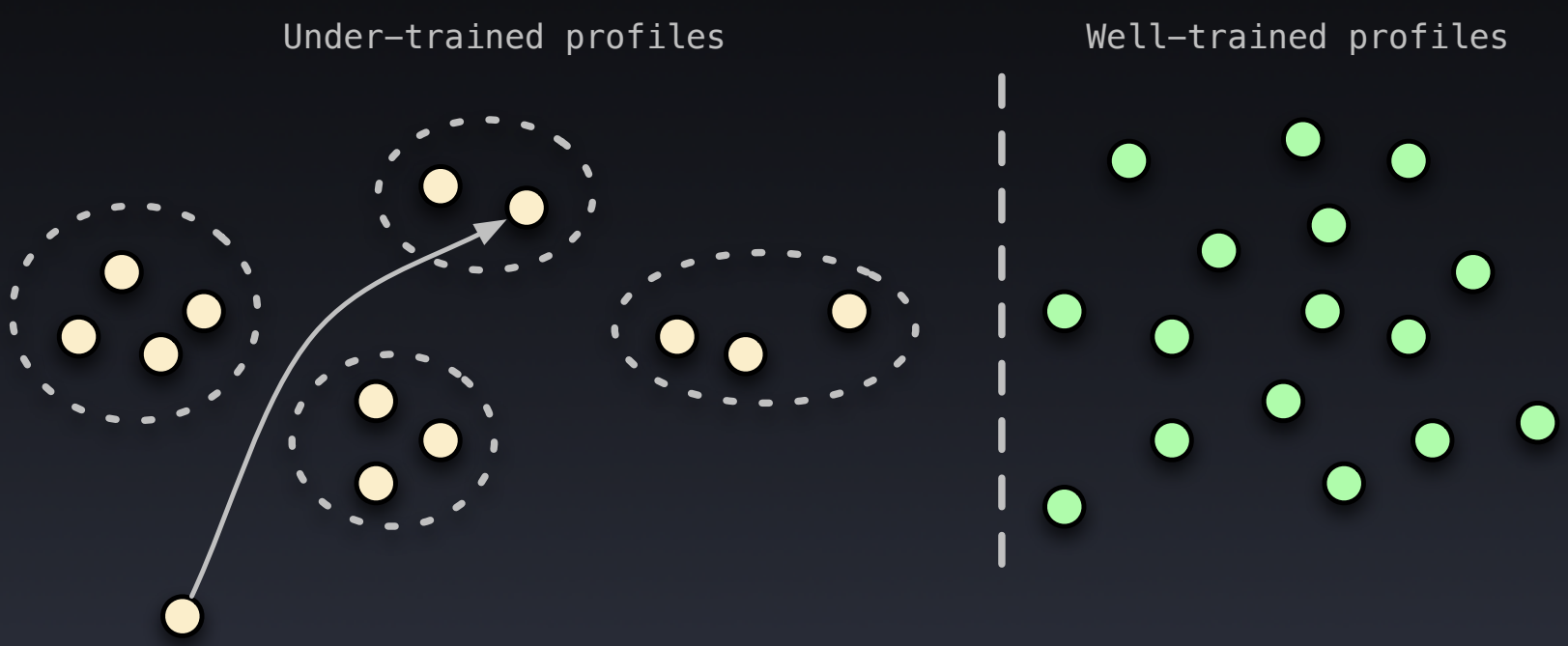




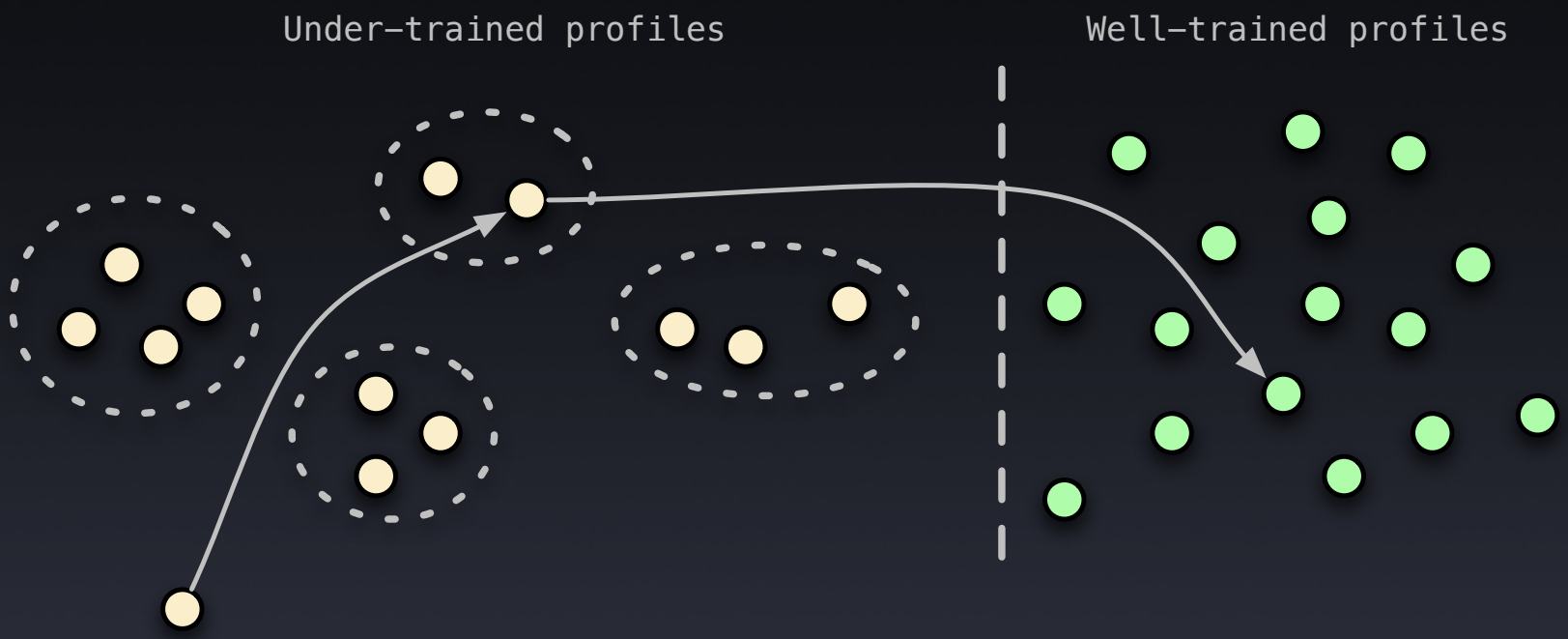
Substituting under-trained profiles



Substituting under-trained profiles



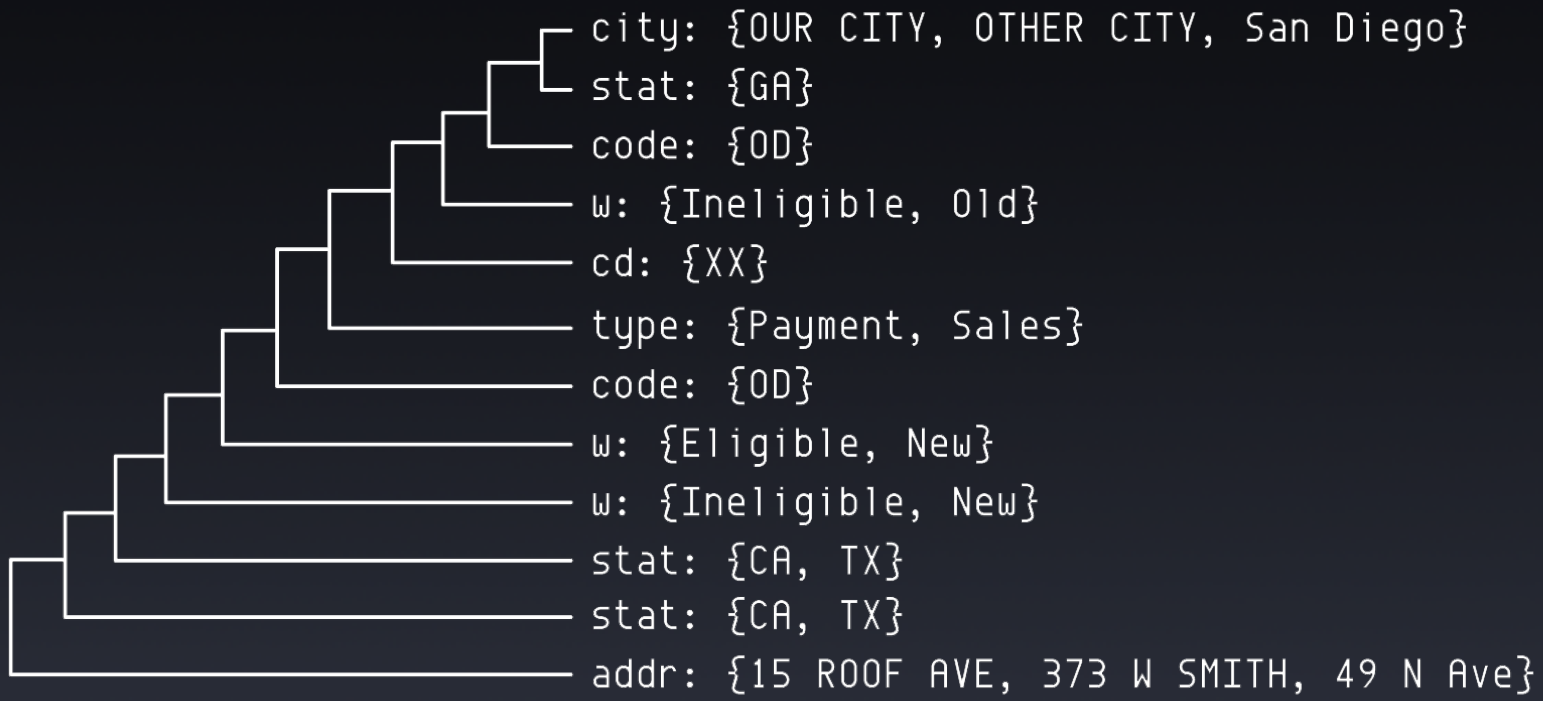
Substituting under-trained profiles



Data set

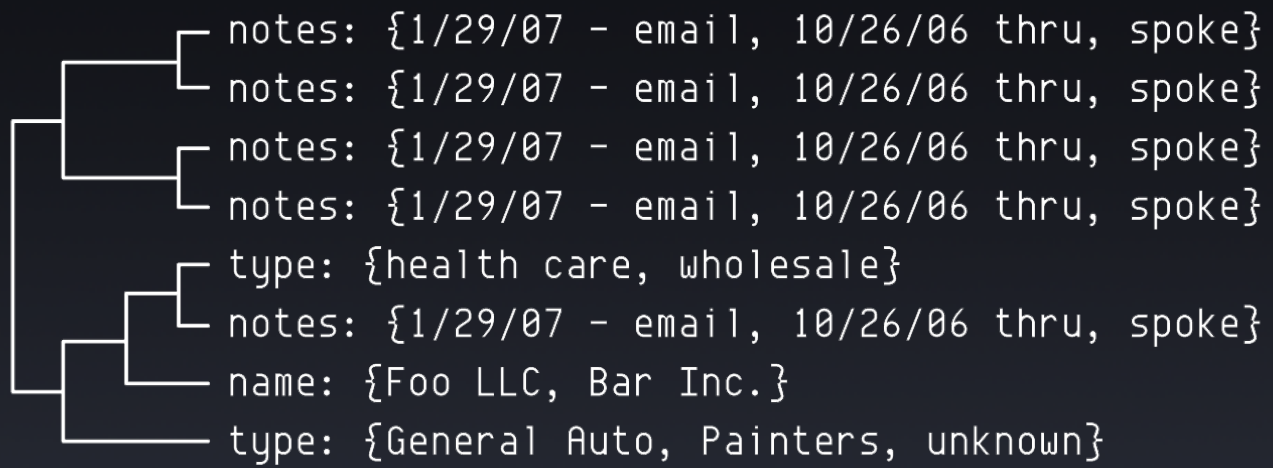
- Real-world web applications from academic, industry domains
- Full content of HTTP connections over 3 months
- 58 million HTTP requests
  - 36,392 unique resources
  - 16,671 unique parameters

Profile clustering quality



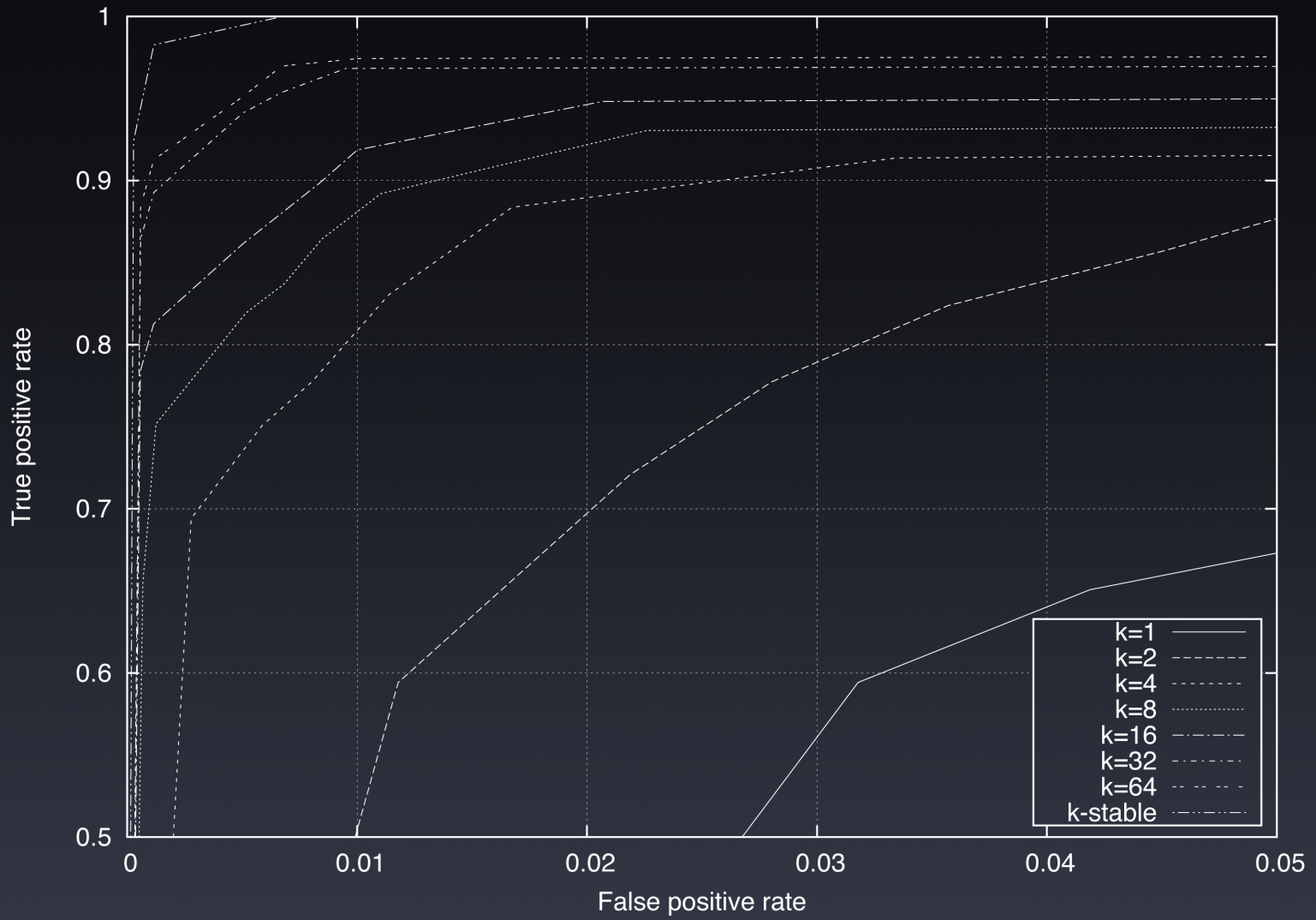
$$\kappa = 8$$

Profile clustering quality



$$\kappa = 64$$

Detection accuracy



Conclusions

- Anomaly detection can suffer from incomplete training data
- But, many features have similar characteristics
- Framework exploits feature similarity to associate under-trained models with well-trained models
- Enhanced sensor shown to perform well over large data set despite incomplete training data