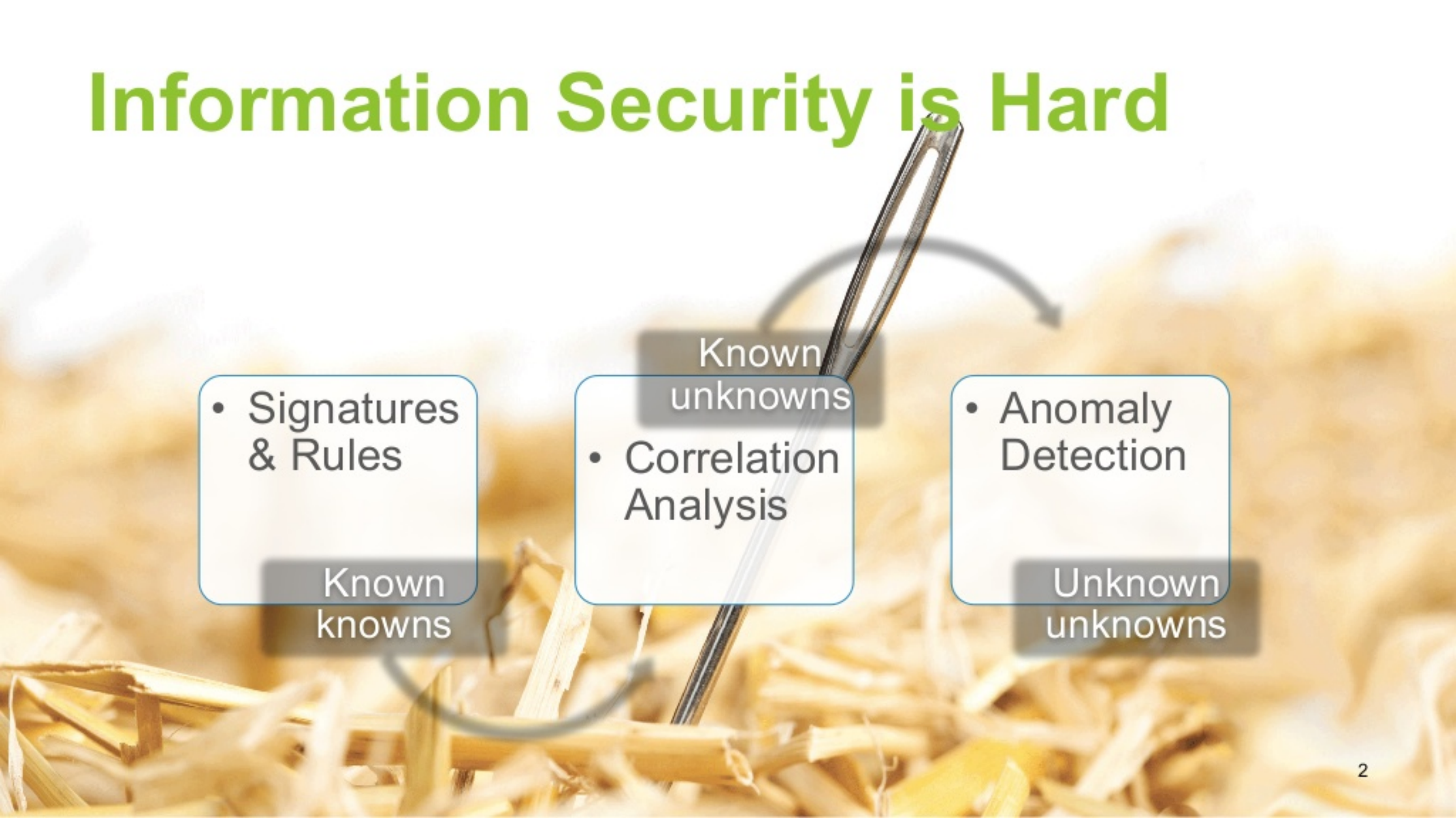
Needle in the Haystack

User Behavior Anomaly Detection for Platform Security

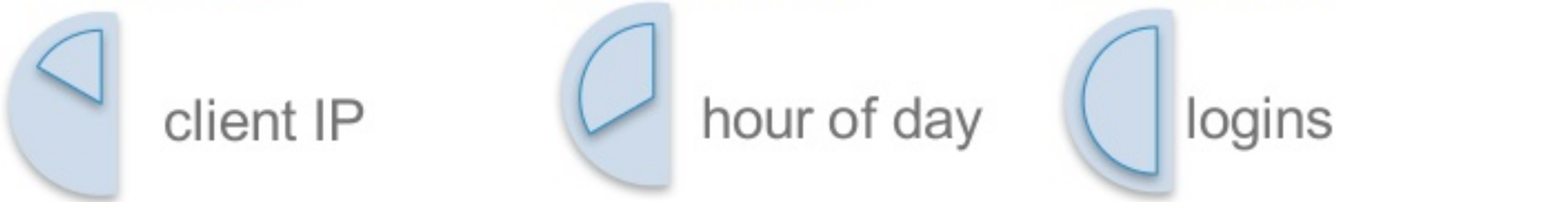
Wei Deng, Ping Yan

Information Security is Hard



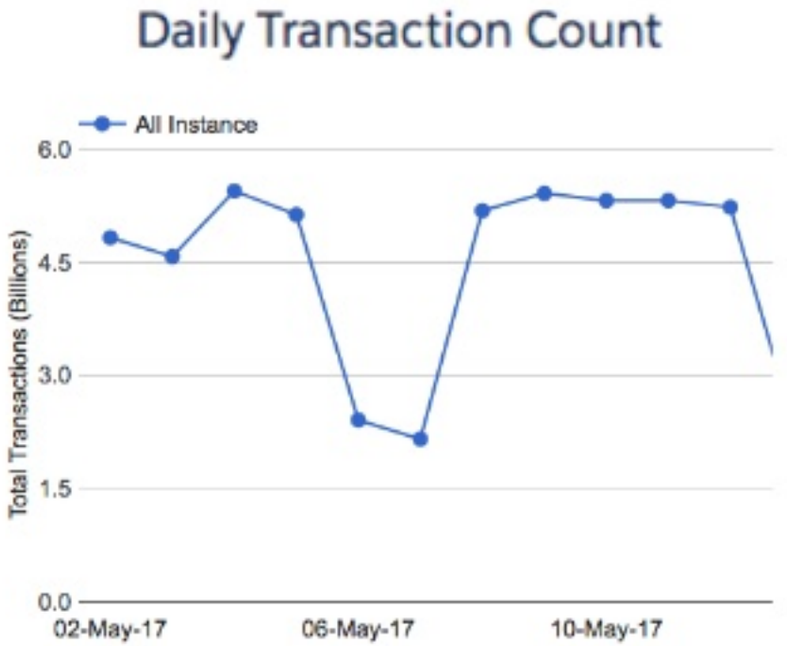
User In-app Behavior

	Entities	Derived	Actions
WHO	client IP	hour of day	logins
WHEN	timestamps	day of week	UI page views
WHERE	user agents	geo country	API calls
HOW	...	...	password reset
WHAT			...

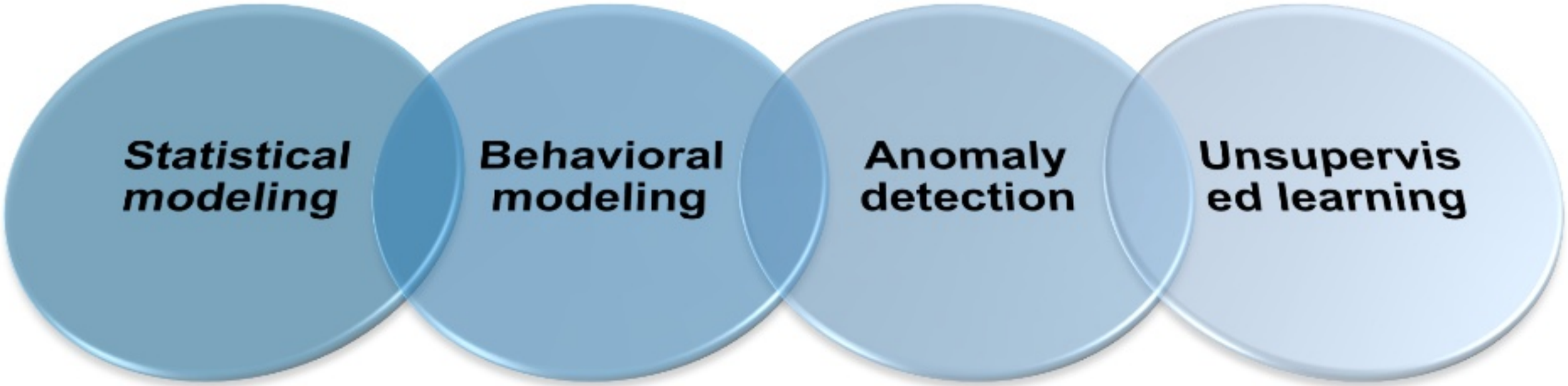


Challenges

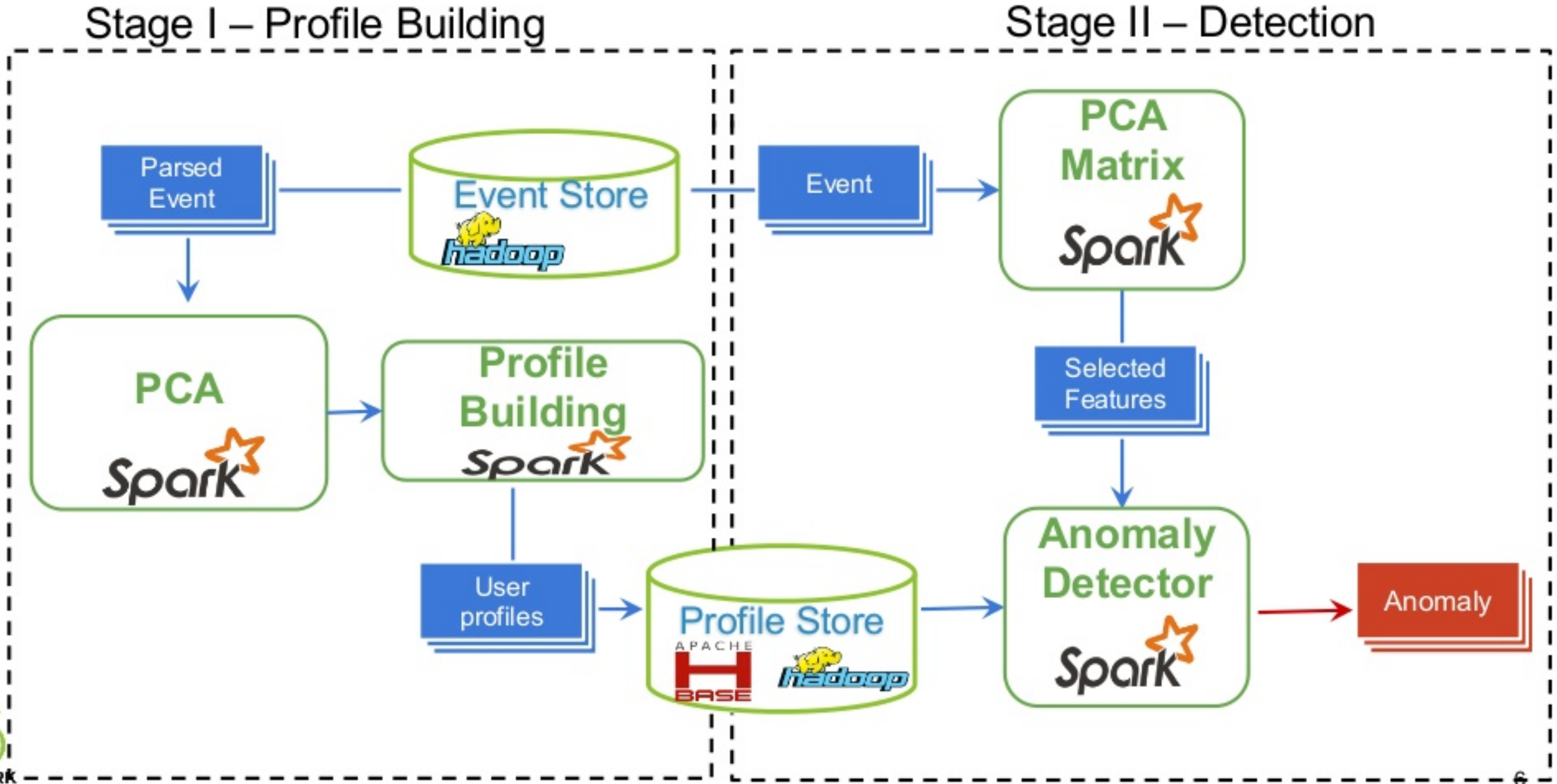
- Size of data, speed to response
- Variability of threats
- Little to none ground truth
- Feature selection
- User behavior novelty



A Behavior Anomaly Detection Approach



Anomaly Detection Engine



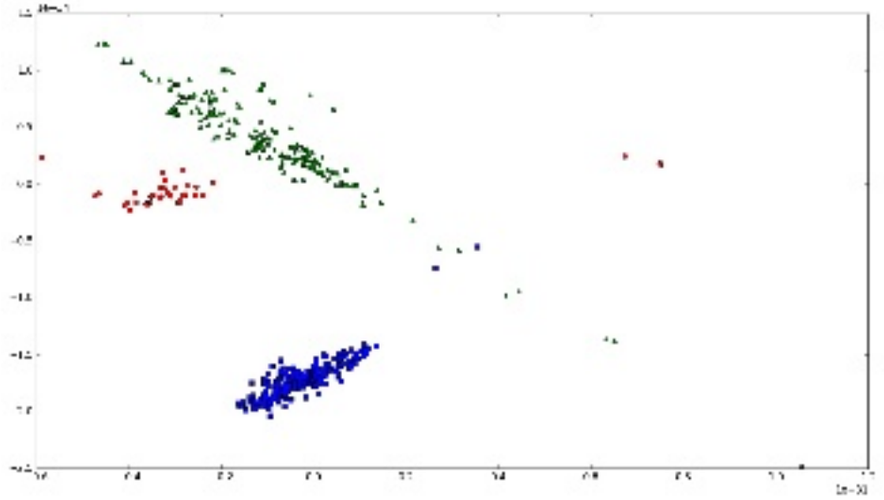
PCA

- High variance variables

    Representative of original data set

- Low variance variables

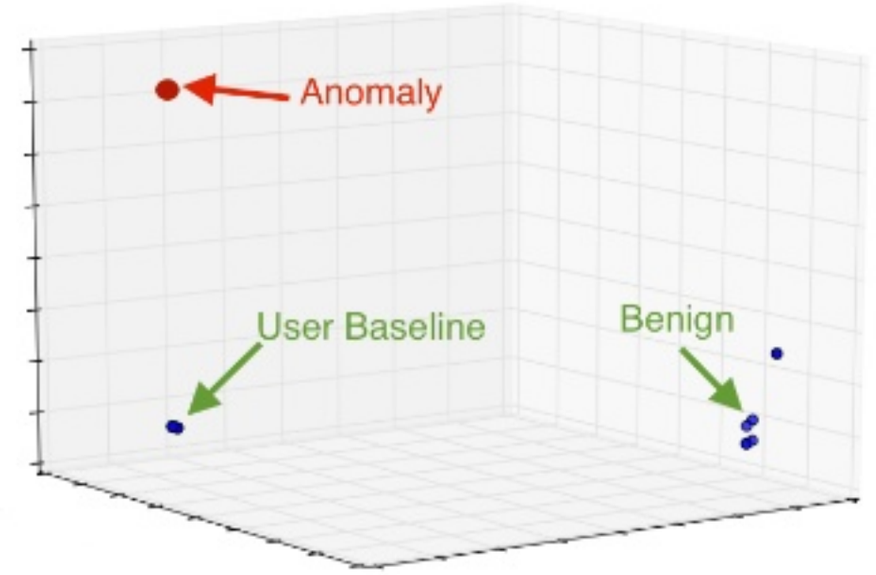
    Representative of users' stable behavior



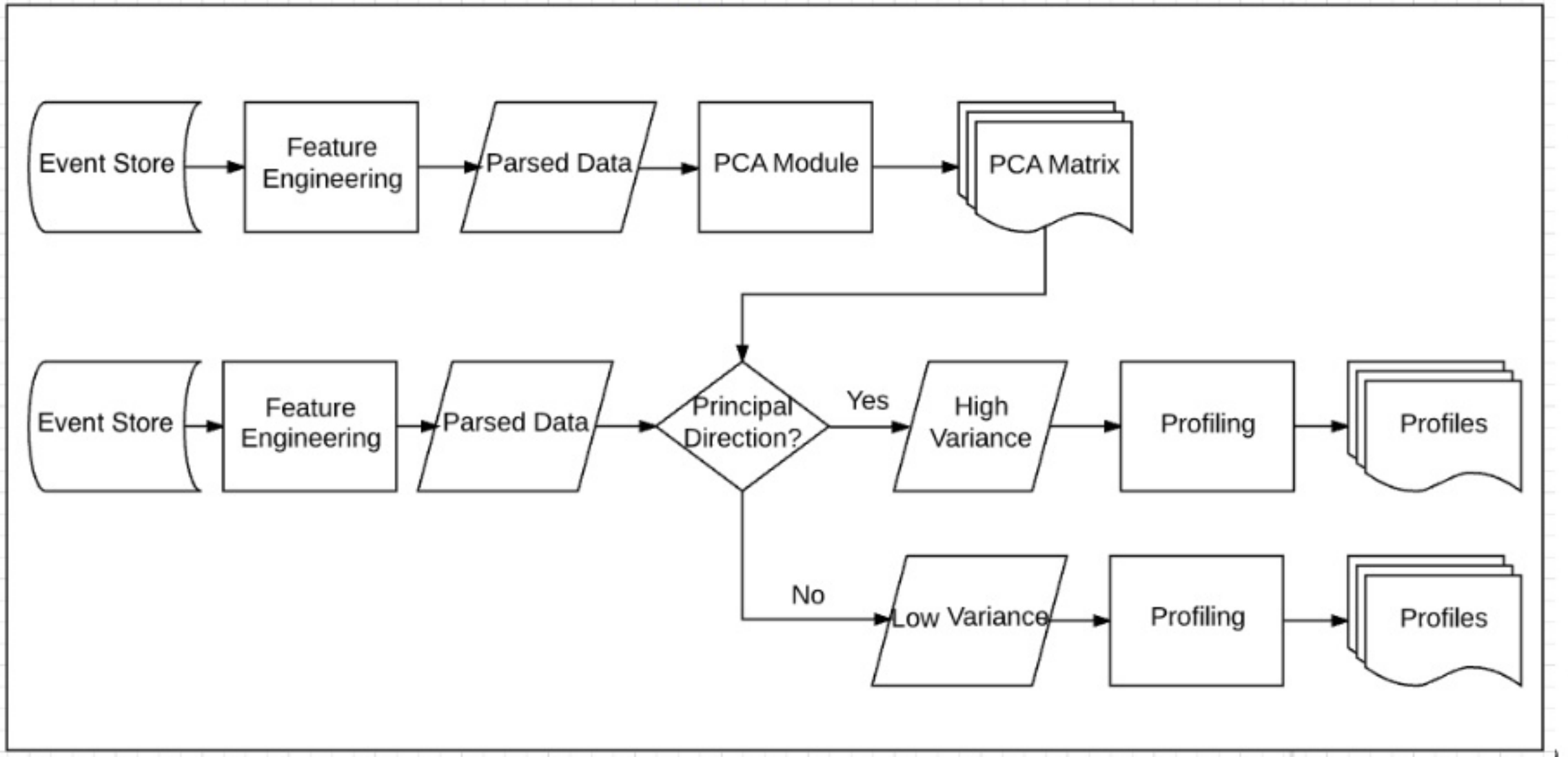
Profile Building

3 statistics to summarize each user's historical distribution

- User's behavior baseline
- Variance of user's behavior
- Legitimate non-typical behavior

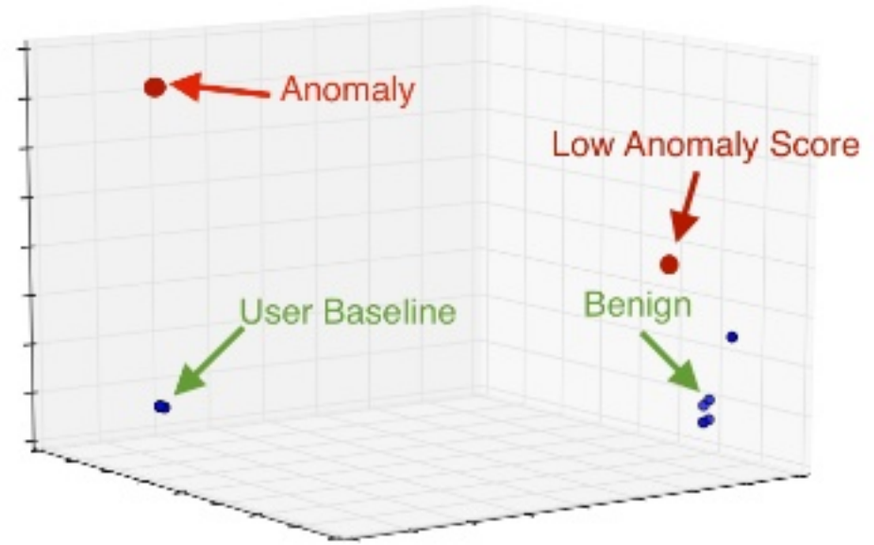


Profile Building

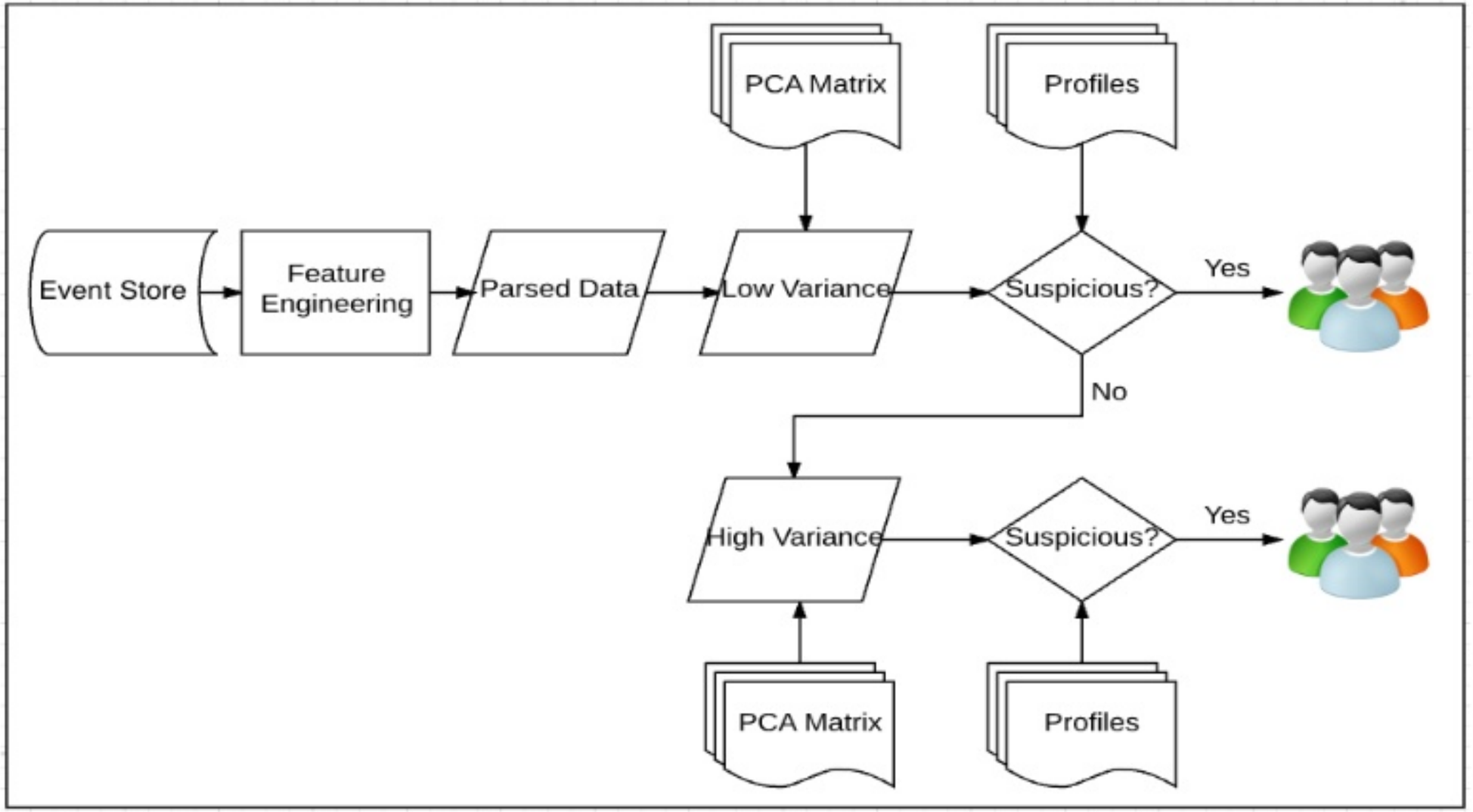


Detection

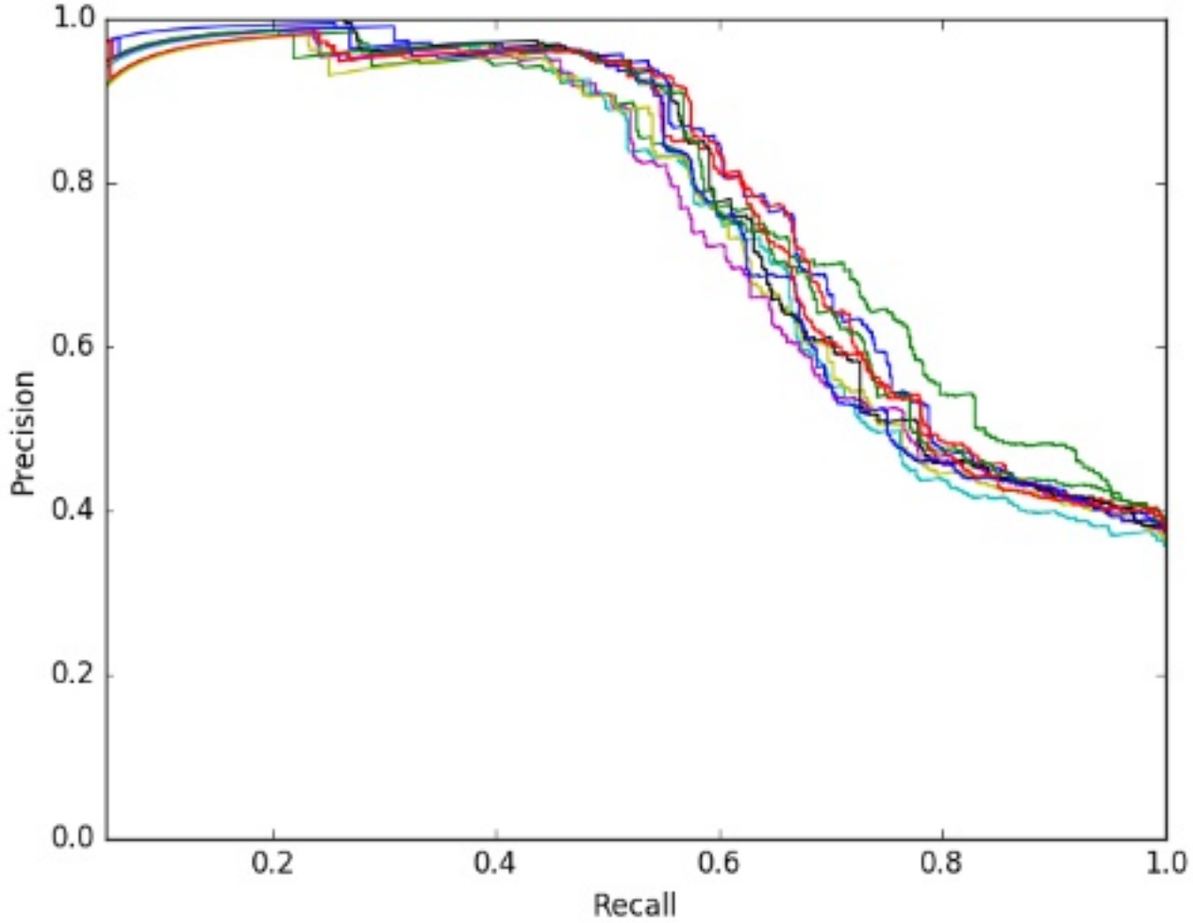
- Deviation from user's baseline
- Re-scale deviation score with a correction factor



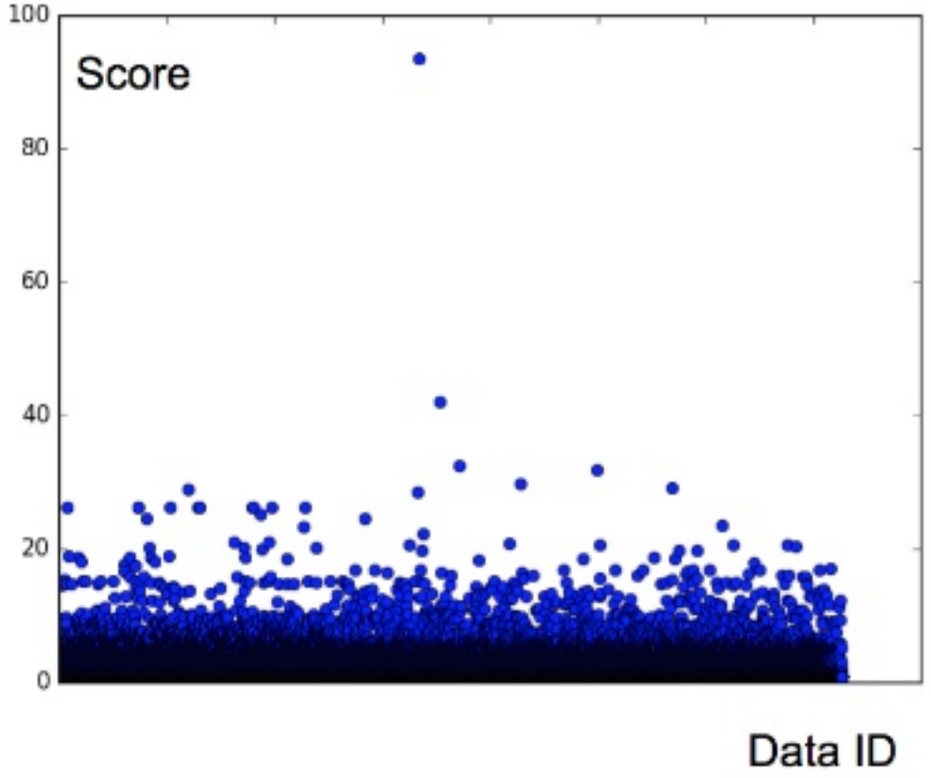
Detection

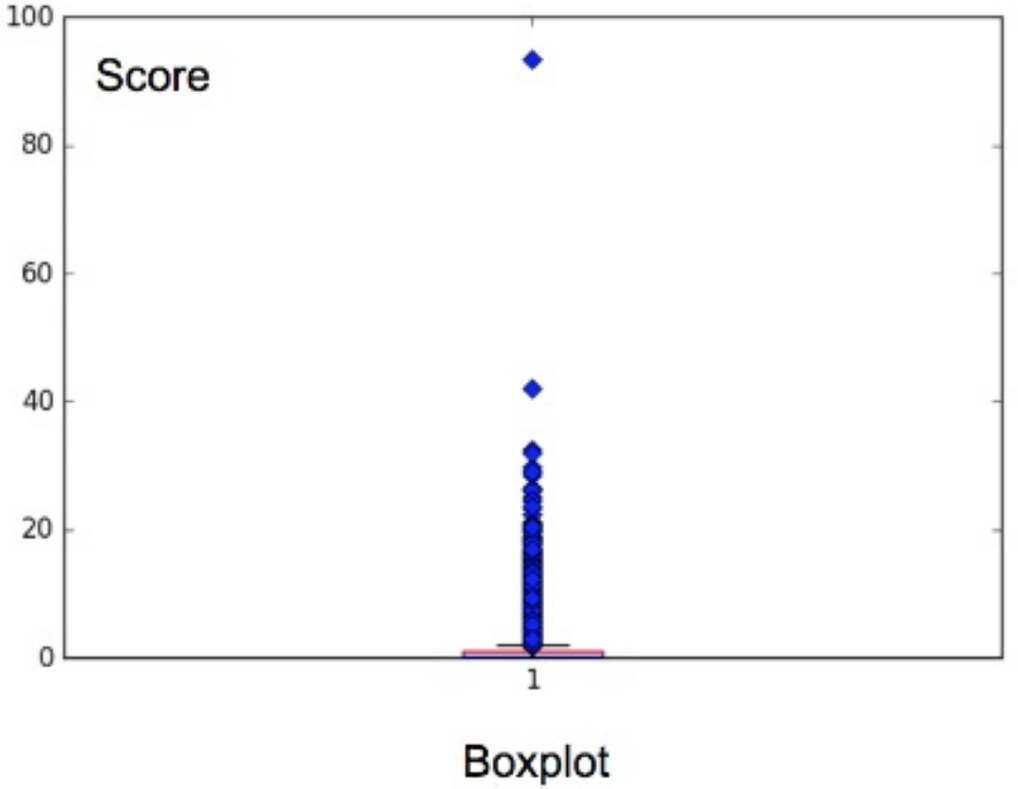

PCA Matrix
Profiles
Event Store
Feature Engineering
Parsed Data
Low Variance
Suspicious?
Yes
No
High Variance
Suspicious?
Yes
PCA Matrix
Profiles

Evaluation with Synthetic Data



Deployment





Thank You.

Wei Deng
wdeng@salesforce.com

Ping Yan
pyan@salesforce.com
@pingpingya