TEL AVIV UNIVERSITY אוניברסיטת תל-אביב *

UNIVERSITY OF CALIFORNIA SANTA CRUZ **

Rudolf: Interactive Rule Refinement System for Fraud Detection

Tova Milo* **Slava Novgorodov*** **Wang-Chiew Tan****

Motivation

- Credit card frauds are unauthorized transactions (billions $ industry)
- Ideally, automatic techniques (ML) should detect and prevent all fraud
- In practice, detection is improved by hand-crafted rules, written by experts that use domain knowledge. One of the main challenges is maintaining the rules and update to prevent new and unknown attacks.

- **Our goal:** guide and assist domain experts in this challenging task.
- **Our solution**: 1. Automatically determine the "best" adaptation to existing rules
  2. Interact with the experts in order to refine the rules, and find the better modification based on their experience and domain knowledge

Work Flow

EXPERT:

- Manually maintaining a lot of rules, searching and adapting relevant rules in order to fix misclassified transactions

RUDOLF:

Given a set of misclassified labeled transactions:

- All misclassified transactions are clustered

Iteratively, on each cluster:

- Find top-k closest rules
- Expert can generalize rules (by extending condition boundaries)
- Experts can specialize rules (via splitting)

Stop when all misclassified transactions are fixed or expert finished

Example

Transaction List:

Time	Amount	Transaction Type	Location	
18:02	107	Online, no CCV	Online Store	FRAUD
18:03	106	Online, no CCV	Online Store	FRAUD
18:04	112	Online, with CCV	Online Store	
19:08	114	Online, no CCV	Online Store	FRAUD
19:10	117	Online, with CCV	Online Store	
20:53	46	Offline, without PIN	GAS Station B	FRAUD
20:54	48	Offline, without PIN	GAS Station B	FRAUD
20:55	44	Offline, without PIN	GAS Station B	FRAUD
20:58	47	Offline, with PIN	Supermarket	
21:01	49	Offline, with PIN	GAS Station A	
:	:	:	:	:

Existing fraud rules Φ on day n:

1) Time $\in$ [18:00, 18:05] $\wedge$ Amt $\geq$ 110
2) Time $\in$ [18:55, 19:00] $\wedge$ Amt $\geq$ 110
3) Time $\in$ [21:00, 21:15] $\wedge$ Amt $\geq$ 40 $\wedge$ Location$\leq$'Gas Station A'

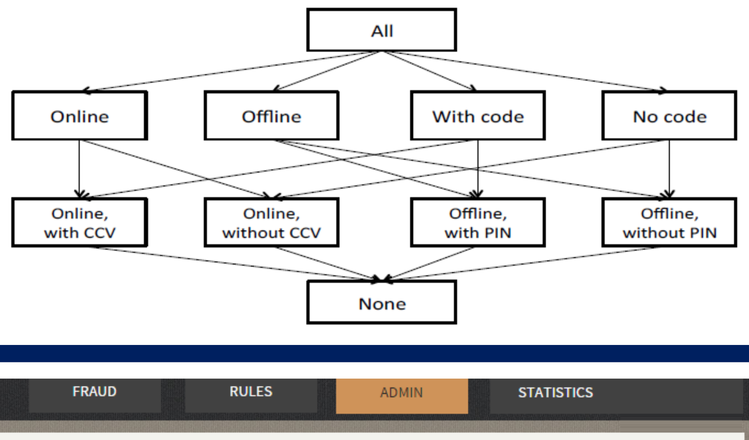
RUDOLF in action:

Time	Amount	Transaction Type	Location
Clustering:			
[18:02, 18:03]	[106, 107]	Online, no CCV	Online Store
Closest Rule:			
[18:00, 18:05]	[110,]	*	*
Generalization:			
[18:00, 18:05]	[106,]	*	*
Specialization:			
[18:00, 18:05]	[106,]	Online, no CCV	*
Expert interaction:			
[18:00, 18:05]	[105,]	Online, no CCV	*

Attribute's Types

Comparison operator:

1. **Numerical – natural**
2. **Categorical – using given ontology**



System Architecture