GOLDRUSH: Rule Sharing System for Fraud Detection

Ariel Jarovsky, Tova Milo, Slava Novgorodov — Tel Aviv University

Wang-Chiew Tan — Megagon Labs

Motivation

- Writing rules to capture precisely fraudulent transactions is a challenging task where domain experts spend significant effort and time.
- Typically, such experts work as "**lone rangers**".
- In fact, there is a lot of commonality in what experts are trying to achieve.

Example

Expert A (USA) Transactions:

Time	Amount	Type	Country	Label
15:58	107K	Stock Trade	Dinotopia	L
16:01	104K	Stock Trade	Dinotopia	F
16:02	111K	Stock Trade	Jamonia	F
16:04	102K	Stock Trade	Dinotopia	F
16:15	96K	Stock Trade	Dinotopia	L
:	:	:	:	:

$$\varphi^A: \text{Type} = \text{“}\mathit{Stock\ Trade}\text{”} \wedge \text{Amount} \geq 100K \wedge \text{Time} \geq 16{:}00 \wedge \text{Country} \in \{\mathit{Dinotopia},\ \mathit{Jamonia}\}$$

Expert B (France) Transactions:

Time	Amount	Type	Country	Label
19:53	140K	Stock Trade	Orsinia	L
20:02	97K	Stock Trade	Orsinia	F
20:03	230K	Stock Trade	Orsinia	F
20:05	92K	Stock Trade	Orsinia	L
20:07	206K	Stock Trade	Orsinia	F
:	:	:	:	:

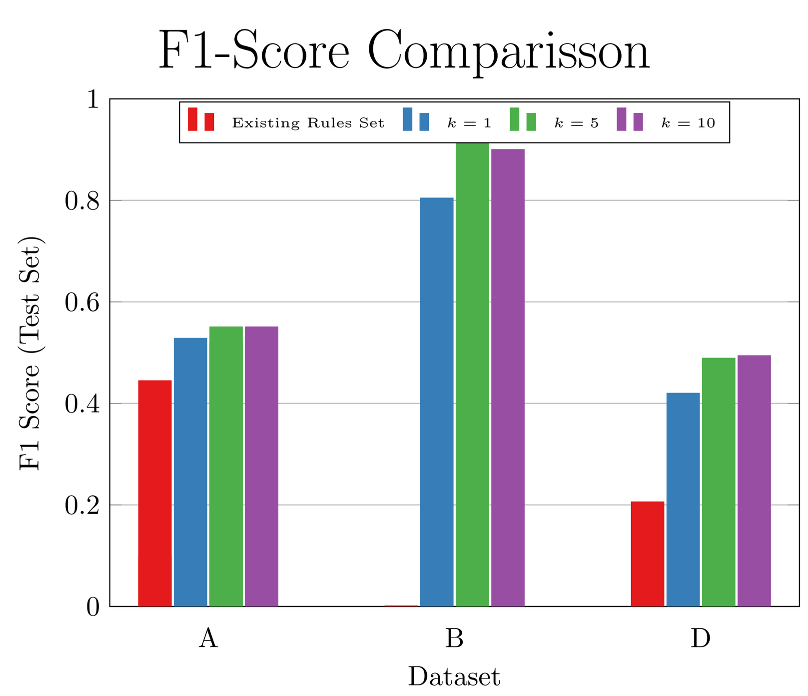
Mapping of $\{Amount \geq 100K\}$ from context A to B:

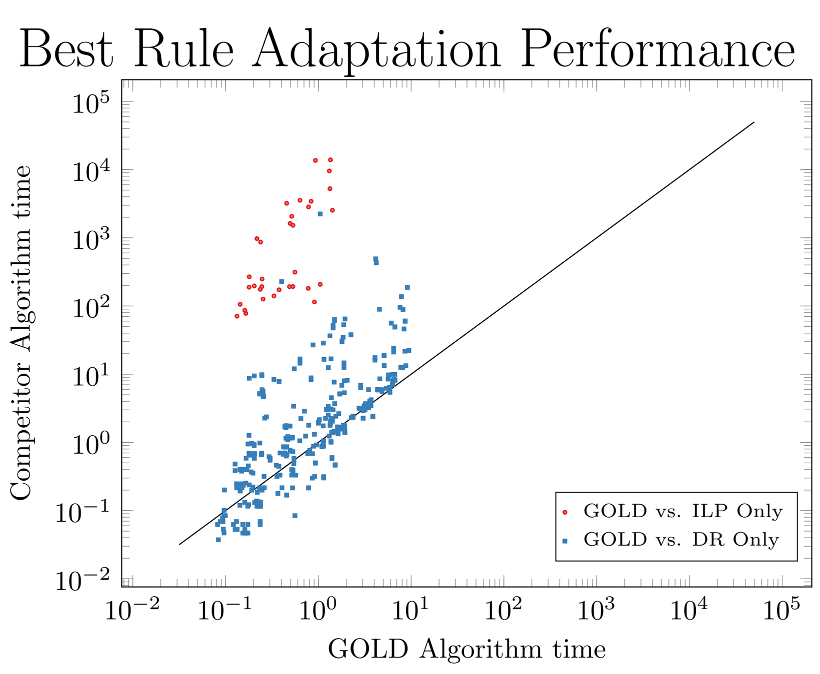
Semantics	Abstraction	Concretization
Identity	100K	100K
Currency Conversion	*97K (CHF)*	95K
Ammounts Distribution	*upper 5%*	200K
Local Regulation Limits	*after hours*	120K

Example of resulting rule adaptation for expert B:

$$\varphi^B: \text{Type} = \text{“}\mathit{Stock\ Trade}\text{”} \wedge \text{Amount} \geq 95K \wedge \text{Time} \geq 20{:}00 \wedge \text{Country} \in \{\mathit{Orsinia}\}$$

Experimental Evaluation





Finding the Best Rule Adaptation

Given a specific rule φ:

- Compute for each attribute its set of semantic mapping candidates: $V_1, \ldots, V_m$
- Compute the set of candidate rule adaptations:

$$\Psi(\varphi) = \{\varphi[v'_1/v_1, \ldots, v'_m/v_m] \mid v'_1 \in V_1, \ldots, v'_m \in V_m\}$$

- Find the rule $\varphi' \in \Psi$ which best improves the current expert's rules set using a linear cost and benefit model:

$$w(\varphi') = (\alpha \cdot |\varphi'(F_C)| + \beta \cdot |\varphi'(F_U)|) - (\gamma \cdot |\varphi'(L_C)| + \delta \cdot |\varphi'(L_U)|)$$

- Evaluation is exponential on the number of attributes (NP-Hard!)
- We build an ILP model of the problem and solve using an ILP Solver

Data Reduction

- The ILP model size is linear on the number of transactions.
- In order to turn it practically efficient even for millions of transactions we developed a Data Reduction technique.
- For example, assume that:
  - $V_{Amount} = \{95K, 100K, 120K, 200K\}$
  - $V_{Time} = \{16{:}00, 20{:}00\}$
- Then the third and the fifth tupples of Expert B will be "**indistinguishable**" no matter which rule will be chosen, and so we can cluster them into a single tuple with a counter:

Time	Amount	Type	Country	Label	Count
20:00	200K	Stock Trade	Orsinia	F	2

- By this way, we can cluster all the "indistinguishable" tuples in the transactions relation, storing a counter for each label:

Time	Amount	Type	Country	F_C	F_U	L_C	L_U
16:00	120K	Stock Trade	Orsinia	0	0	0	1
20:00	$\perp$	Stock Trade	Orsinia	0	0	0	1
20:00	95K	Stock Trade	Orsinia	0	1	0	0
20:00	200K	Stock Trade	Orsinia	0	2	0	0

- Finally, we can solve an adapted ILP model with a smaller transaction relation and with the counters inside the target function.

k-Rules Adaptation

- Generalization of the Best Rule Adaptation Problem for recommending k rule adaptations
- The goal: improve the expert's rule set Fraud Detection accuracy
- Our algorithm uses a prunning technique which, in practice, cuts 66% to 75% of the Best Rule Adaptation algorithm executions.