Detection and Explanation of Anomalous Payment Behavior in Real-Time Gross Settlement Systems

Ron Triepels[1,2]   Hennie Daniels[1,3]   Ronald Heijmans[2]

15th Payment System Simulator Seminar
Helsinki, Finland

31 August - 1 September 2017

[1]Tilburg University, [2]De Nederlandsche Bank, [3]Erasmus University

Introduction

RTGS Systems:

- Facilitate the settlement of financial transactions
- Settle transactions gross and (almost) real-time

Systemic Risk:

> *"The risk associated with any event that threatens the stability of a financial system as a whole" (Berndsen, et al., 2016).*

Research Goal:

- Apply Machine Learning to analyze payment data
- Automatically identify anomalies (stress or undesired behavior)

Anomaly Detection

Anomaly:

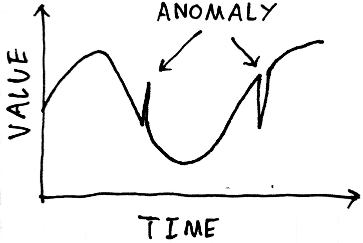
> *"A pattern that does not conform to expected behavior" (Chandola et al., 2009).*

Unsupervised Anomaly Detection:

> *The task of automatically identifying anomalies in a set of unlabeled data.*

Components:

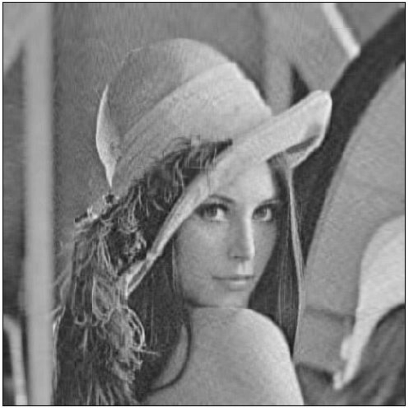
- Model of 'normal' behavior
- Distance function



Lossy Compression

Lossy compression preserves the most important features of data.

Original Picture



Reconstructed Picture

Definitions

Let $\mathcal{B} = \{b_1, \ldots, b_n\}$ be a set of n banks and $\mathcal{T} =< t_1, \ldots, t_m >$ be an ordered set of m time intervals.

We extract $\mathcal{D} = \{\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(m)}\}$ a set of m liquidity matrices from a RTGS system where each $\mathbf{A}^{(k)} \in \mathcal{D}$ is:

$$\mathbf{A}^{(k)} = \begin{bmatrix} a_{11}^{(k)} & \cdots & a_{1n}^{(k)} \\ \vdots & \ddots & \vdots \\ a_{n1}^{(k)} & \cdots & a_{nn}^{(k)} \end{bmatrix} \tag{1}$$

Each element $a_{ij}^{(k)}$ is the liquidity flow between b_i and b_j at t_k.

Liquidity Vector:

$$\mathbf{a}^{(k)} = [a_{11}^{(k)}, \ldots, a_{n1}^{(k)}, \ldots, a_{1n}^{(k)}, \ldots, a_{nn}^{(k)}]^T \tag{2}$$

Anomaly Detection Task

Let $\mathcal{M}$ be a lossy compression model. We measure the reconstruction error of $\mathbf{a}^{(k)}$ after its compressed and reconstructed by $\mathcal{M}$ by:

$$\mathrm{RE}(\mathbf{a}^{(k)}) = \frac{1}{2}||\hat{\mathbf{a}}^{(k)} - \mathbf{a}^{(k)}||_2^2 \tag{3}$$

Accordingly, we classify $\mathbf{a}^{(k)}$ by:

$$h(\mathbf{a}^{(k)}) = \begin{cases} 1 & \text{if } \mathrm{RE}(\mathbf{a}^{(k)}) \geq \epsilon) \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

Here, $\epsilon > 0$ is a threshold.

Autoencoder

We employ a three-layered autoencoder to compress and reconstruct liquidity vectors. The autoencoder can be defined by two functions:

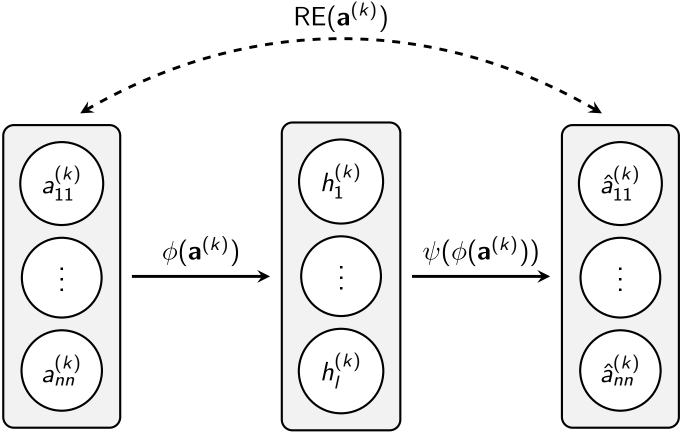
Encoder function ϕ:

$$\phi(\mathbf{a}^{(k)}) = f^{(l)}(\mathbf{W}_1\mathbf{a}^{(k)} + \mathbf{b}_1) \tag{5}$$

Decoder function ψ:

$$\psi(\phi(\mathbf{a}^{(k)})) = g^{(n^2)}(\mathbf{W}_2\phi(\mathbf{a}^{(k)}) + \mathbf{b}_2) \tag{6}$$

Autoencoder Architecture



Model Learning

Parameters $\theta = \{\mathbf{W}_1, \mathbf{W}_2, \mathbf{b}_1, \mathbf{b}_2\}$ are estimated from historic liquidity vectors. We do this by minimizing the following cost function:

$$\mathcal{J}(\theta) = \frac{1}{2m} \sum_{k=1}^{m} ||\psi(\phi(\mathbf{a}^{(k)})) - \mathbf{a}^{(k)}||_2^2 + \frac{\lambda}{2} \sum_{i=1}^{2} ||\mathbf{W}_i||_F^2 \tag{7}$$

Here, λ is a regularization parameter.

We apply stochastic gradient descent in conjunction with back-propagation to solve this optimization problem: I.o.w an optimization algorithm.

Experimental Setup

Payment Data:

- 2.3 million client payments from TARGET2-NL
- Jan 2014 - Oct 2015
- Aggregated over 4,680 consecutive hours
- 20 largest banks

Two autoencoders:

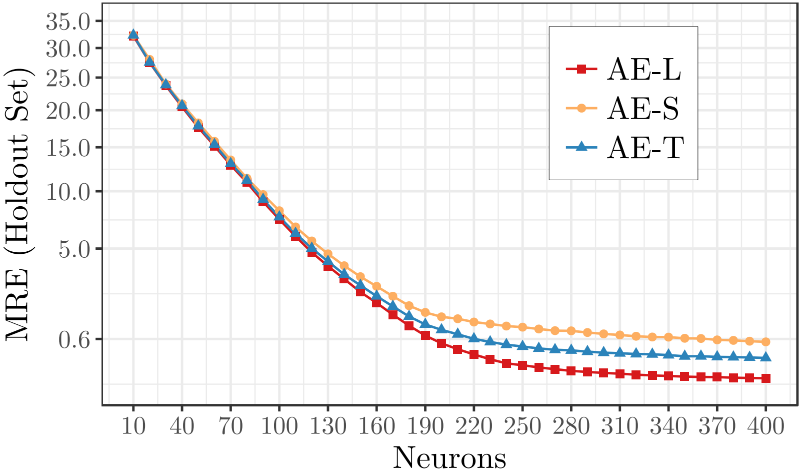
- Linear (AE-L) with (linear/linear) activations
- Non-linear (AE-S) with (sigmoid/linear) activations

Data partitioning:

- Holdout set (2 months)
- Training set (16 months)
- Test set (4 months)

Grid search (1/2)

The number of neurons was optimized by a grid search.



Commercial Bank Run Simulation

Choose a bank b_i and increase its outflow to each $b_j \in \mathcal{B}$ over time:

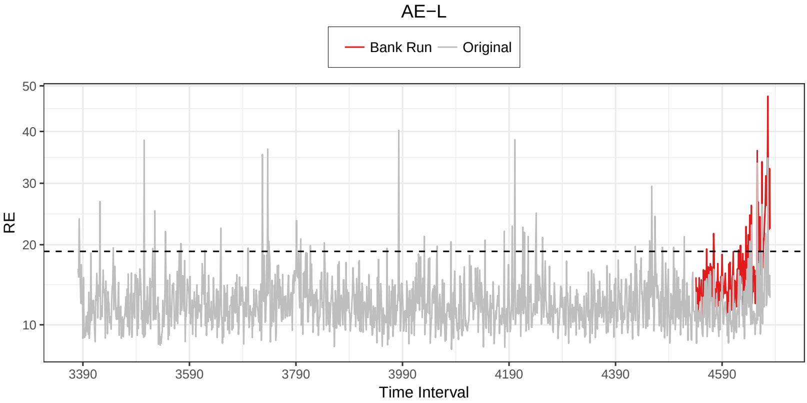
$$a_{ij}^{(k)} := a_{ij}^{(k)} + c_{ij}^{(k)} d_{ij}^{(k)} \tag{8}$$

where:

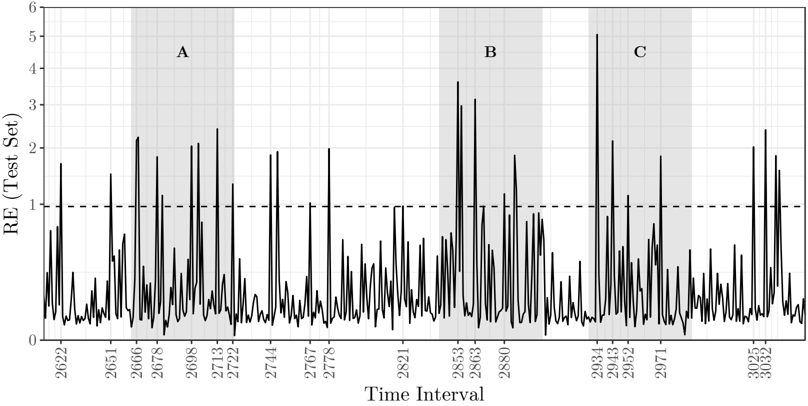
- $c_{ij}^{(k)} \sim B(1, p_{ij}^{(k)})$ determines if liquidity is added.
- $d_{ij}^{(k)} \sim \text{Exp}(\delta_{ij}^{(k)})$ is the amount of additional liquidity.

			Multipliers			
	Rate	**Duration**	p_s	p_e	δ_s	δ_e
Baseline	2	140	1	2	0.1	0.01

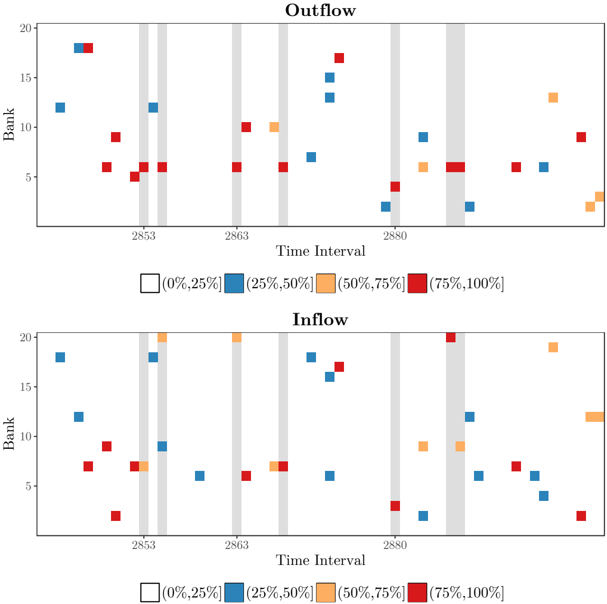
Baseline Simulation of AE-L



Anomalies in real data (1/3)



Anomalies in real data (2/3)



Anomalies in real data (3/3)

Bank run did not follow ‘normal’ pattern because of:

- Initially continuous outflow.
- Part of the ‘gaps’ had no payments (no liquidity, clients no access to accounts).
- Increased flows still considered ‘normal’.

Conclusions

- Autoencoder can detect anomalous flows reasonably well.
- Start of bank run well detected.
- However, part of the anomalous flows during bank run missed.

Questions?