Mining the Czech Insolvency Proceedings Data

Iveta Mrázová and Peter Zvirinský

Department of Theoretical Computer Science and Mathematical Logic

Faculty of Mathematics and Physics, Charles University
Malostranské náměstí 25, 118 00 Praha, Česká republika
Outline

1. Introduction
2. Mining the Czech Insolvency Proceedings Data I (structured data)
   - Insolvency register
   - Data preparation and preliminary data analysis
   - State transition analysis
   - Social network analysis
3. Mining the Czech Insolvency Proceedings Data II (unstructured data)
   - Extraction of missing information
   - Social network analysis: dynamics of the network
   - Reasons for debt
4. Summary, Impact and Further Research
1. Why insolvencies?
Insolvency judge Berka sentenced to 9 years in prison for participating in a criminal organization.

Debtors have found a new trick: They can keep postponing the auction of their assets indefinitely.

Economic survey: Czech insolvencies are ineffective. Creditors retrieve only 3.5% of their claims.
Insolvency filings in the Czech Republic

Number of insolvency filings

Year

2008
2009
2010
2011
2012
2013
2014

0
5000
10000
15000
20000
25000
30000
35000
40000
45000
Introduction

Insolvencies

Insolvency filings in Central Europe

![Graph showing the trend of insolvency filings in Central Europe from 2008 to 2014. The lines represent the Czech Republic, Poland, Slovakia, Slovenia, and Hungary. The graph indicates an overall increase in insolvency filings over the years.]
Introduction

Insolvencies

Insolvency filings in Western Europe
Insolvency filings in the United States
2.1. Mining the Czech Insolvency Proceedings Data (structured data)

Data Preparation
Data Preparation

Preliminary data analysis

Age distribution of debtors

Typical debtor: male in his forties

Time spent in one state affects the following transitions
Data Preparation
Preliminary data analysis

Insolvency filings rates

Uninsolvency filings rates

Unemployment rates

Unemployment (%)

11.4
8.26
5.12
2.2. Mining the Czech Insolvency Proceedings Data

Data Mining
Mining the Insolvency Proceedings Data

Used Methods

- Bayesian networks
- Association rule mining
- Social network analysis
  - prominence and role analysis
  - community discovery

Association rule mining:

<table>
<thead>
<tr>
<th>Ins</th>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Natural Person, Female, Discharge</td>
</tr>
<tr>
<td>2</td>
<td>Company, Bankruptcy Order</td>
</tr>
<tr>
<td>3</td>
<td>Natural Person, Female, Discharge</td>
</tr>
<tr>
<td>4</td>
<td>Natural Person, Male, Bankruptcy Order</td>
</tr>
</tbody>
</table>

$r = \{\text{Natural Person, Female}\} \rightarrow \{\text{Discharge}\}$

support($r$) = $2/4 = 0.5$

confidence($r$) = $2 / 2 = 1.0$
Mining the Insolvency Proceedings Data

State transitions analysis

[Diagram showing state transitions with labels such as 'Unresolved', 'Bankruptcy Order', 'Discharge', 'Reopened', etc.]

- Generally allowed transitions
- Generally not allowed transitions, occurs only in special cases
- Transitions based on explicit Supreme Court decision
- Transitions based on appeal

Transitions that occurred in at least 10% of all insolvency proceedings
Transitions that occurred in 1% to 10% of all insolvency proceedings
Transitions that occurred in 0.1% to 1% of all insolvency proceedings
Transitions that occurred in less than 0.1% of all insolvency proceedings
Mining the Insolvency Proceedings Data

State transitions analysis

Original transition model:
- 14 nodes, 42 directed edges
- most nodes are used very rarely and contain redundant information

Simplified transition model:
- 5 nodes, 12 directed edges
- covers 99% of all insolvencies
State transitions analysis

Bayesian networks (1)

- Discharge is the most used method of resolution
- 41% of all discharges were not successful
- Less than 5% of insolvencies start over once they are finished
State transitions analysis
Association rule mining

different decision-making on bankruptcy in each region:

<table>
<thead>
<tr>
<th>Region</th>
<th>Bankruptcy prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Královéhradecký</td>
<td>87 %</td>
</tr>
<tr>
<td>Pardubický</td>
<td>84 %</td>
</tr>
<tr>
<td>Zlínský</td>
<td>61 %</td>
</tr>
<tr>
<td>Jihomoravský</td>
<td>55 %</td>
</tr>
</tbody>
</table>

Likely causes
- different interpretation of the Insolvency Act
- regional specifics
- abuse of the insolvency law
Mining the Insolvency Proceedings Data

Social network analysis

◎ nodes:
  ○ debtors
  ○ creditors
  ○ insolvency administrators

◎ relations:
  ○ debtor **owes** a creditor
  ○ administrator **manages** the insolvency of a debtor
  ○ debtor **shares** his address with a different debtor
Social network analysis

Ústecký region

- creditors (1000)
- debtors (4500)
- administrators (200)
Social network analysis

GE Money Bank

◎ most frequent creditor
○ participates in 20% of all (27 000) insolvency proceedings

August 2013

GE Money Bank in the Czech Republic is allegedly for sale, the company leaves it without any comment.

The American Concern GE Capital supposedly wants to get rid of the Czech, Hungarian and Polish branches of GE Money Bank. The local bank refused to comment on this rumor.

January 2014

GE Money Bank laid off 70 employees. Tens of job positions are moved from Prague to other cities.

The two months notice period of 70 GE Money Bank employees expires on the end of February. Besides tens of January layoffs the bank also withdrew 16 open positions. The cutbacks involve administration as well as management positions and are a part of the bank’s long term plan.

April 2015

GE Money Bank je na prodej, bude stát desítky miliard

Social network analysis
Community discovery

◎ most frequent creditors:
  ○ GE Money Bank
  ○ Ceská sporitelna
  ○ Cetelem CR
  ○ Provident Financial
  ○ Profi Credit Czech
  ○ Smart Capital
  ○ Home Credit

◎ large groups of debtors
SA1 and SA2 sharing the same address
Social network analysis

Community discovery - SA1

address: Most, Radnicní 1/2, PSC 434 01

(City Hall)
3.1

Mining the Czech Insolvency Proceedings Data II (unstructured data)

Extraction of Missing Information
Extraction of Missing Information

Extracting missing historical data (2)

- determining missing creditors
  - text classification problem
  - Logistic Regression with 96% accuracy
Extraction of Missing Information

Extracting missing historical data (3)

- classifier really learns useful features
Mining the Czech Insolvency Proceedings Data II (unstructed data)

Social network analysis: dynamics of the network (1)
Social network analysis: dynamics of the network (2)
Mining the Czech Insolvency Proceedings Data II (unstructured data)
Dynamics of the network (Senates)

Ústecký kraj
Mining the Czech Insolvency Proceedings Data II (unstructured data)
Dynamics of the network (Senates)

◎ Jihomoravský kraj
Mining the Czech Insolvency Proceedings Data II (unstructured data)

Reasons for Debt (1)

An example of stated debt origin

<table>
<thead>
<tr>
<th>06 Důvod vzniku:</th>
<th>na základě ústní objednávek byly provedeny opravy vozidel a zapůjčení náhradního vozidla</th>
</tr>
</thead>
<tbody>
<tr>
<td>07 Další okolnosti:</td>
<td>faktura byla částečně uhrazena ve výši 26.450,- Kč</td>
</tr>
</tbody>
</table>

Cluster 4

Cluster 5

Cluster 8
Mining the Czech Insolvency Proceedings Data II (unstructed data)

Reasons for Debt (2)
Mining the Czech Insolvency Proceedings Data II (unstructured data)

Reasons for Debt (3)

<table>
<thead>
<tr>
<th>Cluster (Size)</th>
<th>Description</th>
<th>Keywords</th>
<th>Debts per person (in %)</th>
<th>Creditors (in %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Credit card debts and loans</td>
<td>bank product, service product, flexi-account, contract, provision of banking from date, credit, credit card amount</td>
<td>![Graph for Cluster 2]</td>
<td>![Bar chart for Cluster 2]</td>
</tr>
<tr>
<td>3</td>
<td>Credit card debts and loans</td>
<td>contractual, issuing, credit card, revolving, revolving credit, unpaid</td>
<td>![Graph for Cluster 3]</td>
<td>![Bar chart for Cluster 3]</td>
</tr>
<tr>
<td>4</td>
<td>Phone bills</td>
<td>contract, fine, default, phone bill, period</td>
<td>![Graph for Cluster 4]</td>
<td>![Bar chart for Cluster 4]</td>
</tr>
<tr>
<td>5</td>
<td>Seizures and utility bills</td>
<td>seizure, court order, monetary fee, electricity, delivery</td>
<td>![Graph for Cluster 5]</td>
<td>![Bar chart for Cluster 5]</td>
</tr>
<tr>
<td>8</td>
<td>Health insurance bills</td>
<td>insurance, public health, health insurance, premium, penalty</td>
<td>![Graph for Cluster 8]</td>
<td>![Bar chart for Cluster 8]</td>
</tr>
<tr>
<td>14</td>
<td>Credit card debts and loans</td>
<td>credit card, card loan, usage of contract</td>
<td>![Graph for Cluster 14]</td>
<td>![Bar chart for Cluster 14]</td>
</tr>
<tr>
<td>15</td>
<td>Credit card debts and loans</td>
<td>savings account, loan, card loan, limit installment</td>
<td>![Graph for Cluster 15]</td>
<td>![Bar chart for Cluster 15]</td>
</tr>
</tbody>
</table>
4. Summary and Impact
Summary and Impact:
GE Money Bank

GE Money: Insolvenční zákon je dobrý, jen je občas problém se správcí

Část textu

• Každá banka bude mít jinak nastavená pravidla, za kterých pohledávku do insolvenčního řízení přihláší.
• Největší počet pohledávek při osobních bankrotech přihlašuje GE Money Bank, jen loni jich bylo přes 13 tisic.
• Problémem je, že insolvenční správci je hodně a neexistuje cílené řízení ze shora.
Summary and Impact:

Procedural errors (1)
Summary and Impact:
Procedural errors (2)

Robert Pelikán ke zneužívání insolvencí: Všichni vědí, odkud jde dým

- Ať nastavíme zákon jakkoliv, nezabráníme zneužívání insolvencí, říká nový ministr spravedlnosti Robert Pelikán.
- Varuje před častou změnou insolvenčního zákona.
- Kromě toho ale musí řešit i problémy s účelovými, šikanózními či konkurenčními návrhy na insolvenci, kterými se firmy například vyrážejí z boje o veřejné zakázky.
- Speciál o insolvenčích čtěte celý týden v novinách i na webu.

Ať nastavíme zákon jakkoliv, nezabráníme zneužívání insolvencí, říká nový ministr spravedlnosti Robert Pelikán. Lidé z brány podle něj znají jména "podezřelých", ale nejsou důkazy, jejichž pomocí by je bylo možno usvědčit.
3.2

Further Research
Further research

Extract more information and open data
Thank you for your attention!