Data Mining in Financial Markets and an Open Enterprise Infrastructure Supporting Trading & Mining

Chengqi Zhang, Longbing Cao, Jiaqi Wang, Li Lin, Jiarui Ni, Wanli Chen

E-Intelligence Group at FIT, UTS
Data Mining Program at Capital Market CRC
F-TRADE: http://datamining.it.uts.edu.au:8080/tsap
E-Intelligence Group: http://datamining.it.uts.edu.au
Centre of Excellence in Data Mining: http://coedm.it.uts.edu.au
Some problems from financial markets

- Online accessing real huge capital data crossing markets
- Evaluating trading and mining strategies from industry & research in financial markets
- Stock & rules (parameters) association, selection, recommendation, optimization, and integration
- Pattern discovery of (large) market/limit orders in stock markets
- Cross market analysis
- Applications as investment decision support
Overview of the Data Mining Program

Datamining Program
UTS: Broadway/CMCRC: City F-Trade Infrastructure
(integrating DataSources, trading/mining algorithms, offers personalized services in terms of system, data and algorithms)

Infrastructure
Longbing & Jiarui’s work

CMCRC/Tricom/ABN: CBD
(Industrial requirements;
Users: Anybody, anytime, anywhere, from KDD & Finance;
Services: System, algorithms, data)

Industry
Brokers, retailers
Applications like Wanli’s work

Researchers
(Data mining, financial researchers, financial analysis, decision support analysis…)

Data miners
Jiaqi & Li’s work

Team Leader: Prof Chengqi Zhang

Multi-agent & Data-mining
ITR&D-Enabled Capital Market

Australian Technology Park:
Redfern; FIT,UTS
(Diff. Providers: AC3, HK market, ABN, CSFB, etc.
Diff. Formats: FAV, ODBC, JDBC, OLEDB, etc.)

Data & resources
• **Data Mining in** Financial Markets

• **Infrastructure supporting** Trading & Data Mining
Data Mining in Financial Markets
Overview

• Optimization of trading rules
• Correlation analysis in/cross markets
• Pattern analysis of large orders
• Cross market mining
• Platform supporting mining in financial markets
Optimization of trading rules

- Parameter optimization of trading rules
- Rule optimization of technical trading strategies
Optimization of trading rules
Optimization of trading rules

• Moving average
  – constraints..
  – when spot is above the s/r and l/r ma
  – when spot is above the s/r
  – when spot is above the l/r ma
  – when s/r and l/r ma etc etc

• Generate a new optimal moving average
Optimization of trading rules
Correlation analysis in/cross markets

• Correlation between securities
• Correlation between securities and trading rules
• Cross market correlation mining
Correlation analysis in/cross markets

Stock Selection

Market = ASX
RunDate = 2002-06-10
Alg. =

Select  Reset

Stock Selection

The following stocks are recommended according to your strategy:

1. ANP.AK
2. CFU.AK
3. FGL.AK
4. GPT.AK
5. OIT.AK
6. OEC.AK
7. SGB.AK
8. TAH.AK
Correlation analysis in/cross markets

- Pairs trading
  - CBA/GMF
Pattern analysis of large orders

- Pattern analysis of market order and limit order
- Pattern discovery of large orders
Pattern analysis of large orders

January

February

MarketOrder

LargeMarketOrder
Infrastructure supporting Trading & Data Mining
Overview

• What is F-Trade?
• Objectives to build F-Trade
• Benefits for different parties
• System architecture of F-Trade
• Functionalities of F-Trade
• Screenshots @ F-Trade
• Conclusions and future work
What is F-Trade?

• F-Trade is an automated enterprise (agent services-based) infrastructure supporting trading and mining currently on capital markets.
• Online connection to huge, historical, realistic stock information from anywhere.
• Online plug and play, and optimization of algorithms.
• Service provider based on subscription
Objectives to Build F-Trade

- As a trading system (virtual broker)
  - Trading services for data, rules, optimization, evaluation, and reporting
- As a data mining system
  - Mining services for data gateway and preparation, algorithms evaluation, and reporting
- As a virtual service provider
  - Services of system, algorithms, data for mining or trading
Overview

• What is F-Trade?
• Objectives to build F-Trade
• Benefits for different parties
• System architecture of F-Trade
• Functionalities of F-Trade
• Screenshots @ F-Trade
• Conclusions and future work
Benefit for Individual Investors

- **Back-testing**: Individual investors can choose any built-in trading algorithm (or strategy) for any stock they choose to evaluate the performance of their strategy in history.

- **Parameter optimization**: recommend combinations of rule parameters for the best performance of a trading strategy.

- **Signal alerts**: recommend signals for traders
Benefit for Brokers

- **Plug and play & privacy**: Brokers can build their own trading algorithms to plug-in the platform for their private use to evaluate the performance.
- **Huge data**: Broker can benefit from the huge historic real stock information under the platform.
- **Recommendation**: Recommend interesting watchlist of securities, rules, or pairs between securities and rules cross markets.
- **Cross-market**: Brokers can evaluate their trading rules by crossing international markets.
Benefit for Data Miners

- **Data gateway service**: access and extract heterogeneous data sources transparently and remotely if required
- **Data transformer**: transform data as required by some data mining purpose
- **Data stream processing**: data stream preprocessing, mining, and post-processing algorithms with online connectivity to huge amount of stock order stream
- **Multiple data source mining**:
Benefit for Data Miners

- **Privacy-preserving data mining**: Data security-enhancing techniques
- **Integration of data mining**: with database systems, data warehouse systems, and reporting systems
- **Plug-and-play**: Plug-and-play of mining algorithms
- **Evaluation**: back-test/evaluate the plugged mining algorithms
- **Privacy**: keep privacy of the data and the algorithms
Overview

- What is F-Trade?
- Objectives to build F-Trade
- Benefits for different parties
- System architecture of F-Trade
- Functionalities of F-Trade
- Screenshots @ F-Trade
- Conclusions and future work
The System Hierarchy

Users from different fields, at different levels, for different requests may benefit from the system.

Algorithms for different fields can be registered to the system and evaluated.

Multiple programming language supported.

Seamless integration of DS, algorithm, security and components. Offering personalized services to users/programs.

Data warehouse, databases, knowledge base...

Data extraction, cleaning, transformation, loading.

User specific data sources can be plugged into the system.
Technologies underneath F-Trade
F-Trade: the Story So Far

- Administration Center
- Algorithms Center
- Control Center
- Services Center
- User Center
Overview

• What is F-Trade?
• Objectives to build F-Trade
• Benefits for different parties
• System architecture of F-Trade
• Functionalities of F-Trade
• Screenshots @ F-Trade
• Conclusions and future work
Functionalities of F-Trade

• **Data services:**
  – Data source connections
  – Data plug-in
  – Data preprocessing and post-processing

• **Algorithm services:**
  – Algorithm plug and play
  – Trading rules, mining algorithms, etc.
  – Evaluation (back-testing), optimization, recommendation

• **System services**
  – Agent service-based module plug and play
  – Online accessing
  – Automated interface configuration
  – System personalization and reconstruction

• **Trading/mining service provider:**
  – Subscription-based services utilization
  – Huge stock data
  – Trading/mining system
Functionalities of F-Trade

• **Transparent**: users, data sources, algorithms, services, operational systems
• **Flexible**: plug in, reconstruction, interfaces, user profiles
• **Safe**: your privacy of data and algorithms can be kept
• **Online**: registration, plug in, subscription, and iterative evaluation of algorithms, data sources, components
• **Distributed**: platform, data sources, users
• **Profitable**: High opportunities for commercialization
• **Customizable**: F-Trade can be personalized, customized or re-constructed online on your demand
Overview

- What is F-Trade?
- Objectives to build F-Trade
- Benefits for different parties
- System architecture of F-Trade
- Functionalities of F-Trade
- Screenshots @ F-Trade
- Conclusions and future work
### Financial Trading Rules Automated Development & Evaluation (F-TRADE 2.0)

<table>
<thead>
<tr>
<th>Name</th>
<th>Algorithm Library</th>
<th>Trade Strategy</th>
<th>Start Date</th>
<th>End Date</th>
<th>Inputs/Output/Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAC &amp; Forex Flow Basic</td>
<td>MA &amp; MAC Basic 1 1 1 1</td>
<td>Trend Analysis</td>
<td>2003-10-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAC &amp; Forex Flow Enhanced</td>
<td>MA &amp; MAC Basic 1 1 1 1</td>
<td>Trend Analysis</td>
<td>2003-07-15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAC &amp; Forex Flow Optimized II</td>
<td>MA &amp; MAC Basic 1 1 1 1</td>
<td>Trend Analysis</td>
<td>2003-07-15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Inputs/Output/Equation**
  - **Inputs**
    - Spread cost
    - Sharpe ratio
  - **Output**
    - Trend analysis
  - **Equation**
    - MAC & Forex Flow

---

**Notes:**
- The algorithms listed above are part of the F-TRADE 2.0 platform, which automatically develops and evaluates financial trading rules using various strategies and data sets.
- The platform supports multiple inputs and outputs, allowing for comprehensive analysis and optimization of trading strategies.
- Users can customize the trade strategies and adjust parameters for enhanced accuracy and performance.
Intelligent System Reconstruction: Function Tree

Financial Trading Rules Automated Development & Evaluation (F-TRADE 2.0)
Intelligent System Reconstruction: User Permission

UserType Permission Control

UserType List

<table>
<thead>
<tr>
<th>UserType ID</th>
<th>UserType Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>000001</td>
<td>subscriber</td>
</tr>
<tr>
<td>000002</td>
<td>userdirector</td>
</tr>
<tr>
<td>000003</td>
<td>public</td>
</tr>
<tr>
<td>000004</td>
<td>ccce</td>
</tr>
<tr>
<td>000005</td>
<td>administrator</td>
</tr>
<tr>
<td>000006</td>
<td>algprovider</td>
</tr>
</tbody>
</table>
Plug in & Intelligent Interface
Configuration of Algorithms
Iterative Running, Testing, Evaluation of Algorithms
Online services-based usage of algorithms
Summary & Future Work

• F-Trade as an infrastructure supporting trading and mining in stock markets
• Data mining in stock markets through the infrastructure

• As a trading system
  – Semi-language for coding trading strategies in the F-Trade
  – Trading services
• As a data mining system
  – Mining services
• Data mining in stock markets
  – Stock order stream pre-processing & post-processing (visualization)
  – Optimization of trading strategies using data mining
  – Correlation analysis of stocks/rules/stock-rule pairs
  – Pattern discovery of (large) orders based on order book
  – Risk analysis and management using data mining
  – Cross market mining
• Thank you for your attention!
• Any suggestion is welcome!

• Contact us:
  – Chengqi Zhang          Longbing Cao
    Chengqi@it.uts.edu.au   lbcao@it.uts.edu.au
    0438439029             0423255257
    www-staff.it.uts.edu.au/~chengqi www-staff.it.uts.edu.au/~lbcao