Exploring Large Scale Insider Trading Data: Network Patterns & Discoveries

Acar Tamersoy, Bo Xie, Stephen Lenkey, Bryan Routledge, Polo Chau, Sham Navathe

Georgia Tech GT SEC CMU GT GT
tamersoy@gatech.edu

Summary

How do corporate insiders really trade? Does the CEO of a company trade differently than the CFO? We performed the first academic, large-scale analysis of the full insider trading data from SEC, from 1986 to 2012, totaling more than 12M transactions, among 370K insiders. We found that insiders form tightly-connected clusters in which trade related information might propagate.

Insiders and (Illegal) Insider Trading

Financial regulators are interested in applying data mining techniques to detect illegal trades among insiders (e.g., CEOs, directors), by analyzing their Form 4 filings. We performed the first, large-scale academic study of the complete Form 4 filings from SEC.

Insiders engage in illegal insider trading when they exploit their roles and use nonpublic insider information to profitably trade for their companies’ stock.

Form 4 Dataset

SEC requires insiders to disclose their trades within 2 days via Form 4, publicly available from SEC’s EDGAR system (www.sec.gov/edgar.shtml).

We analyzed all forms from January 1986 to August 2012.

Insiders: 370,627
Companies: 15,598
Transactions: 12,360,325
Sale transactions: 3,206,175
Purchase transactions: 1,206,038

Each form contains insider’s name + company + role in the company (from CEO to Retired), transaction date and type (we focused on Purchases and Sales), etc.

Patterns, Observations, & Analysis

We conjecture that some insiders share nonpublic insider information with each other. We build insider networks where nodes are insiders and edges connect insiders trading similarly.

Our similarity function takes as input the transaction times of two insiders of the same company and returns a similarity scale based on the transaction timings.

\[ S(X_C, Y_C) = \frac{1}{|X_C| \times |Y_C|} \sum_{i=1}^{|X_C|} \sum_{j=1}^{|Y_C|} I(x_i, y_j) \]

where \( I(x, y) = 1 \) if \( x = y \)
\( I(x, y) = 0 \) otherwise.

We compute a similarity value for each pair of insiders (\( X_C, Y_C \)) of company C. If both insiders traded at least \( h_t \) times and their similarity value is at least \( h_n \), we include nodes and an edge for these insiders to our network.

<table>
<thead>
<tr>
<th>Network</th>
<th>Nodes</th>
<th>Edges</th>
<th>Connected Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sale</td>
<td>1630</td>
<td>1473</td>
<td>623</td>
</tr>
<tr>
<td>Purchase</td>
<td>1678</td>
<td>2656</td>
<td>489</td>
</tr>
</tbody>
</table>

\( h_t = 5 \)
\( h_n = 0.5 \)

Discussion of Case Studies

Insiders from the same family trade similarly, ~7% of the directly connected insiders share the same last names. All insiders in the chain below belong to the same investment firm, who may be acting on behalf of the firm.

Vertical and horizontal information flow between insiders.

The Securities and Exchange Commission, as a matter of policy, disclaims responsibility for any private publication or statement by any of its employees. The views expressed herein are those of the author and do not necessarily reflect the views of the Commission or of the author’s colleagues on the staff of the Commission.