[Research Paper](https://poloclub.github.io/polochau/insider/asonam13_insider.pdf) Key Points

1. Temporal and network-centric analyses can derive information from insider trading: company roles, transaction types, company sectors, relationships between insiders.
2. *What are other factors? Or other high-level analyses?Read relevant research papers.*
3. *How can we do machine learning to understand the data?*
4. This multimodal nature of factors also includes external factors such as market cycles and regulations

*What are these factors? Explore these variables to better predict*

1. Usefulness of temporal analysis: trading behaviours are affected by economic events and regulations. Hence, examine those that exhibit anomalous activities.

*Explore and then train data?*

1. Usefulness of network analysis: illegal insiders share information within social network. Hence, uncover the hidden communication channels.

*Algorithms to better detect networks. Tune the hyperparameters*

1. Found strong evidence that trade-related information propagates vertically and horizontally.

*One of the factors*

1. Insiders can be classified to routine traders and opportunist traders [Cohen et al.]. This paper explores trading behaviours from network-centric perspective.

*If insiders do not trade routinely, and they earn a high percentage (abnormal), then flag. Useful for machine learning. Can use prediction and statistical method 预测+统计方法*

1. Insiders share negative news before buying, and positive news before selling or stock option exercises. This research ignores the events such as conference calls and public news.

*Could we obtain more data regarding a company, and track the time the insider traded? If the time it drops is earlier than the selling time, flag it. Machine learning can determine if the dropping time is indeed earlier, as there is a long period of dropping. To derive a single discrete value for the dropping time, we need machine learning.*

1. Suspicious accounts can be identified by using classic classification methods such as naïve bayes and logistic regression [Sherlock 13]

Visualisations Made (5.61GB of Form 4)

1. Summary statistics

Text

Description automatically generated with medium confidence

1. Number of companies having a particular number of insiders

Chart, line chart, scatter chart

Description automatically generated

1. Time series by transaction type: sales increase when it is more conducive to earn, decreases when market is bad and selling stocks are worthless. Sales are higher than purchase because insiders receive compensation in stocks. They just rebalance their portfolio for diversification and liquidity for consumption.

Black line denotes 180-day centered moving average for sale transactions. Sale transactions decrease after Quant Meltdown possibly because the stock options became worthless as they were issued at-the-money as compensation prior to the market collapse.

Application, Word

Description automatically generated with medium confidence

1. To analyse trends by **identifying factors and evaluate its importance**, use time as x-axis (constant variable), and break down data according to possible factors – transaction types, role codes and company sectors.

Possible Factors

1. **[Factor 1] Time Series by Role code**: Beneficial owners are not actual insiders – CEOs are. This can be seen from the difference in trading behaviours. “Their selling activity is increased only towards the eve of the financial crisis. Shortly after the crisis, their activity level keeps decreasing even though the transactions of other insiders fluctuate during the same period.”

More transactions by officers and directors as they are the majority.

CEOs start selling aggressively after 2003 and stop doing so in late 2007.

Chart

Description automatically generated

1. Chart

   Description automatically generated **[Factor 2] Time Series by Company Sector (based on transactions)**: Technology sector is the largest sector, similar trend as sales trend possibly because tech companies “prefer to compensate executives with equity”.
2. Chart, line chart

   Description automatically generated Line chart

   Description automatically generated with medium confidence **[Factor 3] Transaction Intervals**
   1. Graph 1: P->P is higher than S->S because tech companies give stock grants so liquidation keeps happening. Oscillatory pattern with 90-day cycle due to corporate bylaws that prohibit transactions near quarterly earnings announcements.
   2. Graph 2: Both have oscillatory pattern with 6-month cycle due to Short- Swing Profit Rule, 16(b) of Securities Exchange Act 1934 (profits from S->P and P->S within 6 months must be disgorged)

If an insider trades according to the oscillatory pattern probably it is a routine trade.

1. Chart, histogram

   Description automatically generated

[Factor 4] Transaction Intervals by Role Codes:

* 1. Beneficial owners are effectively outsiders as there is no oscillatory pattern.
  2. Officers receive stock grants so they have higher P->P than S->S.
  3. Directors are fewer than officers and have less stock compensations.

1. Chart, histogram

   Description automatically generated[Factor 5] Transaction Intervals by Insiders’ Companies’ Sectors:
   1. tech firm insiders sell more than they purchase (as expected). For other sectors, sales and purchases are more balanced.
   2. Still can observe the oscillatory pattern

Network-Centric Analysis

1. Network-Centric Algorithms: things to note

Text, letter

Description automatically generated

* 1. Compare company by company (line 2) instead of all possible pairs of insiders.
  2. Number of distinct (hence use set: avoid split transactions to affect the result) transactions must cross a threshold (so that if the number of transactions is small number, the result will not be affected).
  3. Similar transactions are determined as transactions that occur on the same date. *Machine learning to check if there is a sweet range. Perhaps consider a time window (recommended by author) but the price suddenly fluctuates before or after the event, then flag)*

Text

Description automatically generated with medium confidence

* 1. Similarity function must cross a threshold before it the two insiders can be considered as nodes. *Get another similarity function (recommended by author)*
  2. 7% of the directly connected insiders in the networks share the same last names
  3. Only add the nodes and edges if they previously are not added.
  4. Build graph separately for sale and purchase because they have different implications.

The connected components are within the firms as this is still firm-by-firm analysis.

Text

Description automatically generated with low confidence

A picture containing text, watch

Description automatically generated

1. Size of Connected Component: Chart

   Description automatically generated

This shows that transaction times can discriminate insiders and extract patterns.

1. Chart, waterfall chart

   Description automatically generated

Components with more than 2 insiders are more often tightly connected as the clustering coefficient (mean of local clustering coefficient) is high (bars on the right). This indicates information flow. *Flag such insiders*.

1. Whether information flows outside one company: mostly within a company itself

Table

Description automatically generated

1. Vertical and Horizontal Information Flow

Graphical user interface

Description automatically generated with medium confidence

* 1. CEOs are linked to many OOs (vertical), and Ds, OOs are linked among themselves respectively (horizontal).
  2. The four role codes are mapped using “R. Edelson and S. Whisenant, “A study of companies with abnormally favorable patterns of executive stock option grant timing,” Social Science Research Network Working Paper Series, 2009.”

1. Persistence of similar transactions: Look at the first and last common transactions between insiders, and classify the difference in days into the four bins.

Chart, bar chart

Description automatically generated

* 1. CEOs are less persistent in purchases than in sales.
  2. Only few insiders would trade similarly for more than 3000 days. *Flag them*