Intention-Disguised Algorithmic Trading
(short paper)

William Yuen[1], Paul Syverson[2], Zhenming Liu[1], and Christopher Thorpe[1]

[1] Harvard University {yuen, zliu, cat}@seas.harvard.edu
[2] Naval Research Laboratory syverson@itd.nrl.navy.mil

1 Introduction

Large market participants (LMPs) must often execute trades while keeping their intentions secret. Sometimes secrecy is required before trades are completed to prevent other traders from anticipating (and exploiting) the price impact of their trades. This is known as "front-running". In other cases, LMPs with proprietary trading strategies wish to keep their positions secret even after trading because their strategies and positions contain valuable information. LMPs include hedge funds, mutual funds, and other specialized market players.

However order information is leaked, or why it is sought, traders who exploit others' order information extract value from markets at the expense of the LMP. Thus, hedge funds and other firms take great pains to hide their intentions, even generating "noise" trades to hide their intended positions from other traders [2]. We present trading schemes that disguise an LMP's intentions and positions from *any* other entity, including the brokers that the LMP interacts with.

Various studies [13, 12] have shown abnormal price behavior and significant negative price impact from information leakage prior to a block trade execution. Thorpe and Parkes [16, 17] discuss cryptographic and security research on exchanges and how information can be exploited in financial markets. But, existing research generally proposes new infrastructures or protocols, for which adoption is notoriously difficult. We take a simpler approach. Our specific contributions are: (1) to propose a general model underlying the design of trading strategies that leak no information, (2) to study major scenarios in the market and design associated algorithms *that require no changes to the existing trading infrastructure*, and (3) to prove those algorithms leak no information in those scenarios. These algorithms can serve as building blocks for more challenging real-world scenarios beyond our present scope. Though our approach is algorithmic, we are not concerned with volume-weighted algorithmic trading. See [4], [5] and [11] for a review of the literature and for insights into the study of automated trading.

We next discuss existing trading infrastructure, define three types of adversaries, and present ways they can extract information from orders placed by the LMP. In Section 3, we describe the model for information leakage and address the needed properties for an efficient trading strategy. Section 4 introduces different trading strategies that disguise the intention and holdings of the LMP from exploiters. We evaluate their defensive performance against each of the three types of "exploiters". Given the available space, our presentation gives only the

basic ideas. Detailed mathematical explanations of various information leaks, as well as theorem statements and proofs, can be found in the full paper [18].

2 Preliminaries

Existing trading infrastructure and exploiters

Brokers include brokers, dealers, and broker-dealers. *Shares* are units of any security, including equities, bonds, currencies, or derivatives. One can long or short any of the shortable securities represented through brokers. Each transaction is for a nonzero integer number of shares; although LMPs typically trade in increments of at least 100 shares. When a trade is executed, the symbol and quantity is publicly reported by the exchange. Typically, only the broker involved in the trade knows the identity of the LMP and whether the LMP was the buyer or seller. From weakest to strongest, the categories of exploiters are:

1. *Curious Observers* are able to see trades printed as they are executed and/or the prices and sizes of orders (requested trades) as they are quoted. With sufficient intelligence and experience, curious observers may be able to guess the identity and intention of the LMP.
2. *Individual Curious Brokers* are able to see trade orders by the LMP before they are executed. A corrupt or careless broker can leak the LMP's intentions for exploitation by broker insiders or external agents. Using multiple brokers, the LMP can limit the information a single curious broker can extract.
3. *Colluding Curious Brokers* are able to see trade orders by the LMP and can share their information with each other. If all brokers used by the LMP are curious and collude, any benefits resulting from splitting trades across different brokers would be lost. However, all brokers used by the LMP must collude in order to yield complete knowledge.

In all our strategies the LMP places a set of orders for an asset d at one or more brokers in order to yield a net purchase or sale of d. Using minimal resources, we want to prevent reasonably capable exploiters from guessing the net order. Strategies must also stay completely effective even when exploiters are aware that the LMP is using them. We focus primarily on the following scenarios:

- *Multiple brokers with one trader* (nB1T): There is only one trader represented in the market, the LMP. This trader can interact with many brokers.
- *Single broker with multiple traders* (1BmT): Only one broker handles trades, for the LMP and possibly for other market participants.
- *Multiple brokers with multiple traders* (nBmT): Multiple traders and multiple brokers can trade simultaneously, the most general scenario.

Two motivating trading strategies and why they leak information

To hide the net order we first consider a simple approach where the LMP uses two brokers and places an order with each one so that neither broker individually learns about the net order. This simple strategy still allows brokers to

Stock	Shares through Broker A	Shares through Broker B	Net LMP volume traded
ATK	+500k	−500k	0
SXL	−300k	+400k	+100k

Table 1. An example of disguising true trading intentions using two brokers.

extract some knowledge. For example in Table 1, Broker A observes that a large ATK trade has gone through Broker B when Broker B prints the block trade after execution. Broker A is not sure whether the LMP is involved in the trade with Broker B, or the direction of the LMP's trade through Broker B. But, A knows that his large client has traded one of three net positions: 500k, 500k + 500k = 1M, or 500k - 500k = 0. Similarly, broker A knows his client's net trade for SXL is 100k, -700k, or -300k shares. Because the number of possible cases is low, Broker A can analyze each scenario and deduce the best exploitation strategy. For example, a block trade price that is closer to the bid than to the offer is more likely to be seller-initiated [10].

Another simple strategy is to use a single broker but multiple registered traders. An LMP might create several registered trading agents that are not known to be associated with the LMP but trade on its behalf. Thus we could produce the exact same structure as Table 1 except that now instead of a single trader using Broker A and Broker B, we have Trader A and Trader B, both responsible for the same book of the LMP, trading through a single broker. The broker cannot tell from what he sees if he is dealing with one LMP shopping two blocks or two LMPs. This defends against collusion, unlike the two-broker system. But, if the broker links the two pseudonymous traders together, then he will know everything about their intentions going forward. We can combine the two solutions so that each pseudonymous trader is splitting orders across multiple brokers. This gains both the advantages and the overhead of both approaches.

3 Defining "information leak"

A rigorous definition of "information leak" is needed to understand both the potential threats from exploitation for the LMP and the desired properties of the trading strategies we are seeking. Here we provide just a sketch of such definitions and refer the reader to the full paper [18]. We propose in this section three types of information leak (or rather its absence) so as to formalize the notion: *zero information leak*, ϵ-*information leak*, and *full space strategy*.

Our inspiration is Goldwasser et al.'s [9] notion *zero knowledge*. Roughly, transmitting a piece of information is zero-knowledge if the universe of computations the recipient can perform does not change after receiving the information.

Definition 1. (Efficient algorithms for zero information leak) *Let* $(\Omega, \mathcal{F}, \Pr)$ *be a probability space that represents all possible intended positions of an LMP and the corresponding a priori distribution over these positions. Let* ω *be a random*

sample from Ω. A trading algorithm $\mathcal{A}$ is said to be perfect-zero-knowledge *with respect to exploiters if the following two conditions hold:*

- *$\mathcal{A}$ can generate an execution plan in polynomial time (wrt a reasonable representation of Ω) that ends with the LMP holding exactly ω shares.*
- *The exploiters are able to generate the distribution on the random variable M on their own without seeing the signal ω.*

A natural relaxation of zero information leak is to allow ϵ information leak. The definition is essentially the same as this except that exploiters can generate a random variable with a statistical difference[1] from M of at most ϵ. One may think of the difference between perfect zero knowledge and statistical zero knowledge [7] to understand the motivation for this relaxation in security definition.

Finally, we propose another way to ensure sufficient noise that an adversary is unable to eliminate any possible values from Ω. Specifically we require that $\Pr[\omega \mid M = p] > 0$ for all q and all ω such that $\Pr[\omega] > 0$.

Definition 2. (Efficient algorithms for full space strategy) *Let $(\Omega, \mathcal{F}, \Pr)$ be a probability space that represents all possible intended positions of an LMP and the corresponding prior over these positions. Wolog, assume that $\Pr[\omega] > 0$ for any ω. Let ω be a random sample from Ω. A trading algorithm $\mathcal{A}$ is said to give a* full space strategy *with respect to exploiters if the following two conditions hold:*

- *$\mathcal{A}$ can generate an execution plan in polynomial time (w.r.t. a reasonable representation of Ω) that ends with the LMP holding exactly ω shares.*
- *For any message M observed by the exploiters, $\Pr[\omega \mid M] > 0$ for any $\omega \in \Omega$.*

Although there are more refined notions of knowledge, e.g., that quantify the exact number of bits leaked by a system [8], it is unclear how the amount of leaked information relates to the financial cost of the information. A single leaked bit information can have great value (the sign of an order issued by an insider), but other times even a large information leak may be harmless.

4 Trading strategies

In this section. we design and analyze trading strategies to counter various adversaries in various markets, and in progressively more challenging scenarios.

Multiple brokers with one trader (nB1T)

In order to defend against the three types of exploiters mentioned, we first build our strategies using a single trader and n orders placed with n different brokers. We call this *the nB1T platform*. We start with nB1T strategies for the LMP against *curious observers* (the weakest). The following sign flipping game is closely related to a trading strategy that leaks no information:

[1] the statistical difference between two discrete random variables X and Y is defined as $\sum_i |\Pr[X = i] - \Pr[Y = i]|$

Definition 3. (Sign Flipping Game) *Given an interval* $[-q, q]$*, find a set of numbers* $T = \{t_1, t_2, ..., t_n\}$ *such that* $\sum_i t_i = q$ *and*
(1) for any integer $x \in [-q, q] \cap \mathbb{Z}$ *there exists a set of numbers* $a_1, a_2, ..., a_n \in \{-1, 1\}$*,* $t_i \in \mathbb{Z}$ *such that* $x = a_1 \cdot t_1 + a_2 \cdot t_2 + ... + a_n \cdot t_n$*,*
(2) The number n *is a function of* q*. The value of* n *should be as small as possible.*

Intuitively, for our nB1T strategy, n in the sign flipping game is the number of brokers the LMP interacts with, and $\Omega = [-q, q]$ is the range of net position the LMP wants to hold. By buying or selling volume t_i with broker i, he can construct every possible desired net trading volume, x, bounded between $-q$ and q. Unsigned traded volumes $T_L = \{|a_i t_i|\}$ are printed among other traded volumes W_0 that do not involve the LMP. Observer identification of T_L from $T_L \cup W_0$ depends on market liquidity and other factors. An LMP is always able to set a larger q at the cost of higher transaction costs. When the security parameter q is fixed, a natural goal is to minimize the number of brokers used.

Now, suppose the LMP wishes to buy $x \in [-q, q]$ shares (negative x notated as selling) of a product. She would then be able to execute a sequence of orders $t_1, t_2, ..., t_n$ to each of the brokers such that $x = t_1 + t_2 + ... + t_n$.

From an observer's point of view, he only sees the sequence $|t_1|, |t_2|, ..., |t_n|$. If he does not have information of the LMP's intention a priori, the observer can only attempt to extract knowledge by going through all combinations of the signs for all t_i. Therefore, the LMP's strategy should make the following set as large as possible: $S = \{a_1|t_1| + a_2|t_2| + ... + a_n|t_n| : a_1, ..., a_n \in \{-1, 1\}\}$. A necessary requirement for a zero-information-leak trading strategy is that $[-q, q] \subseteq S$. Our first goal is to construct $T = \{t_1, t_2, ..., t_n\}$ with minimum possible n such that S fully covers $[-q, q]$. We can find a T with $|T| = \lceil \log_2 q \rceil + 2$ that satisfies the first requirement of the sign flipping game. In fact this is nearly optimal in that any set T that satisfies the first requirement of the sign flipping game will have $|T| \geq \lceil \log_2 q \rceil + 1$. Further, there exist on the nB1T platform both efficient strategies that leak zero information and full space strategies against curious individual brokers. Proofs of these and related results are in the full paper [18].

The above strategies no longer work against *curious individual brokers who do not collude*. For example, in our analysis [18] of efficient strategies for the sign flipping game, curious broker b_n, knowing q and seeing the sign a_n of an order of size $q/2$, would know that the LMP is intending to buy from the range $[-q, 0]$ if $a_n = -1$ or $[1, q]$ if $a_n = 1$. If instead we are less efficient, splitting trades across more brokers, or less complete, making some of the intermediate values unreachable, then we can prevent any one broker from knowing this much about the LMP's position. We will revisit this observation below.

There are also efficient ϵ-information-leak strategies for the curious broker market. When $1/\epsilon$ is a constant or a polynomial in n, the strategy has a $\Omega(\text{poly}(n))$ expansion. See the full paper for rigorous statements and details.

Countering collusion

The above nB1T strategic platform does not yield strong defense against curious colluding brokers: they can share knowledge with each other, including

the identity of the LMP and the sets a_i and t_i. If colluders know the total number of brokers used and can find all of them, the value x can be trivially extracted.

Even if n is not known or not all n brokers collude, certain possible values for x can be eliminated: Suppose, for example, the LMP uses two sets of brokers $R = \{b_1, b_2, ..., b_n\}$ and $R' = \{b'_1, b'_2, ..., b'_n\}$, and that $R \cap R' = \emptyset$. Let $B = R \cup R'$. Suppose brokers $B_c \subset B$ collude and share the information $T_c \subset T$ and $A_c \subset A$ with each other, and let J be the set of indices corresponding to colluding brokers. With enough colluders they can learn significant information. For example, if $\sum_{j \in J} a_j t_j > \sum_{i \notin J} |a_i t_i|$, colluding brokers would know that $1 \leq x \leq q$.

To maximize the collusion resistance for a given n, it is clearly optimal to split q uniformly across all n brokers. In other words, every broker is used to trade q/n shares, either buying or selling. (Let some brokers be allowed to receive no order when n is odd to hide a zero position.) This of course leaks n (easily countered by randomization). Also, note that, even if n is known, the colluding brokers B_c can never learn more than their proportion of the LMP's position.

Single broker with multiple traders (1BmT)

To defend against broker collusion, we now examine utilizing m registered trading agents (hereafter referred to as traders) by the LMP to create the desired net position. The mathematics behind this 1BmT platform is very similar to the nB1T strategy: Simply substitute m traders placing orders at one broker in place of one trader at n brokers (where $m = n$). The same theorems hold for 1BmT as for nB1T. In practice, the additional redundant positions held by the traders add ongoing carrying and transaction costs. Also, changing brokers, especially in a developed market, is generally easier than changing registered traders.

We next consider strategies against the *curious individual broker*, assuming he is unable to identify the traders associated with the LMP. Suppose the LMP places a set of orders $\{a_k t_k\}$ at the broker via m different traders. Let $W_0 = \{w_1, w_2, ..., w_z\}$ be the normal market interest seen by the broker; i.e., the set of orders the broker receives from clients not affiliated with the LMP. The broker thus sees total market interest $W_t = \{a_k t_k\} \cup W_0$. In a very liquid market, $\exists w_i \in W_0 \ni |w_i| = |a_k t_k|$ for $k = 1, ..., m$. In this case, the broker cannot identify any $a_k t_k$ from W_t, and the 1BmT platform does not leak information to him. This is not so when liquidity is low and the broker knows the LMP is employing 1BmT, however. For example, if there is no corresponding surge in activity in the overall market or at other brokers, he can infer that all market interests may originate from the LMP. Furthermore, if $\nexists w_i \in W_0 \ni |w_i| = |a_k t_k|$ for some k, $a_k t_k$ can be identified as originating from the LMP. Thus, elements in the set S can be eliminated, similar to the nB1T platform under collusion. These potential information leaks on the 1BmT platform in an illiquid market motivate our next strategy platform.

Multiple brokers with multiple traders (nBmT)

We can extend the above strategies by using n brokers (with index j) and m traders (with index i), with the security parameter q remaining the same. In general form, the LMP uses the set of traders $\{d_1, d_2, ...d_m\}$, each of the trader d_i places orders with a subset of brokers $\{b_{i1}, b_{i2}, ..., b_{in}\}$. In total, a maximum

of $n \cdot m$ orders are placed with a maximum of $n \cdot m$ unique brokers. In practice, some of the b_{ij}'s are the same broker. One possibility is to split the net order x that the LMP wishes to place into m different orders, $\{a_1t_1, a_2t_2, ...a_mt_m\}$, for m traders as in the sign flipping game in Definition 3. Each trader d_i can then place its individual single order a_it_i, with broker b_{i1}. In this case, the total number of orders placed is $m = \lceil \log_2(q) \rceil$.

Curious observers cannot see the identities of the traders. Thus, nBmT would look the same as nB1T to curious observers. So, as under nB1T, the external observers cannot extract the trade order made by each trader, thus cannot extract any knowledge about the LMP. Another variant (nBmT2) of this strategy is to divide x into m sets of orders $\{a_1t_1, a_2t_2, ...a_mt_m\}$ for m traders according to the sign flipping game. This is detailed in our full paper [18].

We now study the performance of nBmT against *curious individual brokers*. Each trader d_i places its order a_it_i at a different broker. Let W_i be the set of orders each broker b_{i1} receives from his clients not affiliated with the LMP, or normal market activity. Broker b_{i1}, sees total interest $W_{ti} = a_it_i \cup W_i$, and he cannot identify a_it_i from W_i since he he does not know that d_i is affiliated with the LMP. This case is different from 1BmT because the market activities W_v at other brokers $b_{v1}, v \neq i$, are also increasing due to the activity of the LMP in nBmT. Thus, even in a low liquidity environment, broker b_{i1} cannot determine whether the increase in $|W_{ti}|$ is due to the activity of the LMP (the presence of a_it_i), or due to increased general market volume (an increase in $|W_i|$).

Collusion does not benefit brokers if traders $\{d_i\}$ are not revealed to be affiliated with the LMP. Colluding brokers do not know which orders are affiliated with the LMP and therefore would act at worst as a single broker in the 1BmT scenario. Thus, the nBmT strategy can guard against total broker collusion.

5 Conclusions and Future Work

We have examined the problem of placing orders while hiding intention. We presented models of information leakage, and based on these models, we derived three classes of strategies against curious observers, individual curious brokers, and colluding curious brokers.

Though not our current focus, we believe transaction costs of these strategies can sometimes be reasonable, such as when the notional share price is high and/or the bid-offer is tight. We estimate these costs in [18]. We hope this class of intention-disguised algorithmic trading can reduce the profitability of and incentive for exploiting trade information, and alter market behavior as a whole. To that end, understanding these costs, and reducing them, is important.

Open Questions and Future Research

We believe that either finding the lower bound of the brokers that need to be used (in terms of $f(n)$) or finding a better strategy using fewer brokers may be possible. Furthermore, the sign of a trade with any one broker may be inferred by an observer using a trade direction algorithm such as that developed by Ellis, Michaely and O'Hara [6] or Peterson and Sirri [15]. Our strategies are

unaffected, assuming that all trades are filled in one round. However, realistically, such trades may take multiple rounds. On the other hand, in practice, there are also often other market participants trading, thus creating cover noise against identifying the trades initiated by the LMP. Even in an extremely illiquid market with no other active trading participants, the orders being worked by brokers are not synchronous in practice. Therefore, even if a broker with malicious intention is able to deduce the signs of other brokers, he cannot front run confidently that he has seen all the relevant trades initiated by the LMP.

Acknowledgement Zhenming Liu is supported in part by NSF CCF-0634923.

References

1. S. Brain, "A front-running smile?" *Traders Magazine*, May 2005, available online at `http://www.wisdom.weizmann.ac.il/~oded/zk-tut02.html`
2. G. Chacko, personal communication, August 2009.
3. G. Di Crescenzo, "Privacy for the stock market" in *Proc. Financial Cryptography and Data Security*, 2002.
4. I. Domowitz, H. Yegerman, "The cost of algorithmic trading: a first look at comparative performance," in *Algorithmic Trading: Precision, Control, Execution*, March 2005.
5. I. Domowitz, H. Yegerman, "Measuring and interpreting the performance of broker algorithms" in *ITG Inc. Research Report*, August 2005.
6. K. Ellis, R. Michaely, M. O'Hara, "The accuracy of trade classification rule: evidence from NASDAQ," *Journal of Financial and Quantitative Analysis*. 2000.
7. O. Goldreich, "Zero-knowledge: a tutorial." accessed through `http://www.wisdom.weizmann.ac.il/~oded/zk-tut02.html`
8. O. Goldreich, E. Petrank, "Quantifying knowledge complexity," in *32nd IEEE Symposium on Foundations of Computer Science*, 1996.
9. S. Goldwasser, S. Micali, C. Rackoff, "The knowledge complexity of interactive proof systems," in *17th Annual ACM Symposium of Theory of Computing*, 1985.
10. L. Harris, "Trading and exchanges: market microstructure for practitioners," *Oxford University Press*, 2003.
11. M. Kearns, Y. Nevmyvaka, A. Papandreou and K. Sycara, "Electronic Trading in Order-Driven Markets: Efficient Execution", *IEEE Conference on Electronic Commerce (CEC)*, 2005.
12. D. B. Keim, A. Madhavan, "The upstairs market for large-block transactions: analysis and measurement of price effects," *The Review of Financial Studies,* 1996.
13. R. Kumar, A. Sarin, K. Shastri, "The behavior of option Price Around Large Block Transactions in the Underlying Security," in *The Journal of Finance*. 1992.
14. A. Madhavan, "VWAP Strategies," *Investment Guides, Transaction Performance.* Spring 2002.
15. M. Peterson, E. Sirri, "Evaluation of biases in execuation cost estimates using trade and quote data," Forthcoming in *Journal of Financial Markets*. 2002.
16. C. Thorpe, D. C. Parkes, "Cryptographic securities exchanges" in *Financial Cryptography and Data Security*, 2007
17. C. Thorpe, D. C. Parkes, "Cryptographic combinatorial securities exchanges" in *Financial Cryptography and Data Security*, 2009
18. W. Yuen, P. Syverson, Z. Liu, C. Thorpe, "Intention-Disguised Algorithmic Trading," Harvard School of Engineering and Applied Sciences Tech. Report TR-01-10.