

선물시장의 시스템트레이딩에서 동적시간와핑 알고리즘을 이용한 최적매매빈도의 탐색 및 거래전략의 개발

이석준¹ · 오경주²

^{1,2}연세대학교 정보산업공학과

접수 2011년 1월 22일, 수정 2011년 3월 6일, 게재확정 2011년 3월 17일

요약

국내 정치적·사회적·경제적 요인 및 국제 정치 상황, 해외 경제 동향 등의 요인들을 비롯한 IMF 이후의 금융시장 개방에 따른 외국투자자본의 유출입으로 인하여 한국 금융시장의 불확실성은 더욱 증가되었다. 특히 투자자들은 의사결정에 더 많은 혼돈을 겪게 되었고 투자 시 도움을 줄 수 있는 보다 유용한 도구들을 필요로 하게 되었다. 본 연구는 시스템 트레이딩을 이용하여 선물시장에서 거래 할 때 최적의 매매 타이밍을 알아보고 이에 적합한 전략을 알아보는 것이 목적이다. 패턴인식 알고리즘인 동적시간와핑(DTW; Dynamic Time Warping) 알고리즘을 이용하여 빈도별(10분, 30분, 60분, 일별) 유사 패턴을 찾아내고 최적의 매매 타이밍을 분석한다. 이를 위해 주식시장의 대표적인 패턴들을 알아보고, 유사한 패턴을 보이는 기간을 DTW를 이용하여 빈도별로 분석한다. 유사한 패턴들의 검증을 위해 기술적 지표들의 개별 전략을 적용한 거래 시뮬레이션을 실시한다. 시뮬레이션 결과 대부분 30분 데이터에 적용된 전략들이 높은 수익률을 가져왔다.

주요용어: 기술적 분석, 기술적 지표, 동적시간와핑 알고리즘, 선물시장.

1. 서론

국내의 정치적·사회적·경제적 요인 뿐 만아니라 국제 정치 상황, 해외 경제 동향 등의 요인들을 비롯한 IMF 이후의 금융시장 개방에 따른 외국투자자본의 유출입으로 인하여 한국 주식시장의 불확실성은 증가되고 있다(정원길, 2003; 신양규, 2009). 특히 투자자들은 투자의사결정에서 더 많은 혼돈을 겪게 되었다. 따라서 시장에 참여하는 모든 투자자들은 안정적인 수익을 얻을 수 있는 자산운용을 위하여 각종 이론과 시장정보 및 투자자 자신의 경험에 따른 다양한 전략들을 개발하고 있다. 특히, 급격한 정보통신기술의 발달로 인하여 증권시장에 커다란 변화가 왔다. 인터넷의 확산과 고성능 컴퓨터의 보급이 확산됨에 따라 증권시장에는 온라인거래가 활성화 되었고, 개인들은 증권회사 직원의 힘을 빌리지 않고 언제 어디서나 손쉽게 주식 거래를 할 수 있으며 시시각각 변화하는 다양한 금융정보를 얻고 주식 매매 의사결정에 사용하게 되었다. 이러한 현상으로 자연스럽게 스스로 거래 전략을 만들고 적용할 수 있는 시스템 트레이딩의 발전에 따라 투자자들은 객관적이고 안정적인 투자 전략을 구현할 수 있게 되었다.

주식시장 분석에 있어 투자자들은 주가의 패턴에 많은 관심을 가지고 있으며, 나아가 이를 중요한 요소로 생각하고 있다. 주식시장에서의 패턴이란 주가가 움직일 때 나타나는 일정한 형태를 말한다. 예를

¹ (120-749) 서울특별시 서대문구 신촌동 134번지, 연세대학교 정보산업공학과, 박사과정

² 교신저자: (120-749) 서울특별시 서대문구 신촌동 134번지, 연세대학교 정보산업공학과, 부교수.

E-mail: johanoh@yonsei.ac.kr

들어 상승패턴, 하락패턴, 횡보패턴 등이 있다. 최근에 패턴예측을 위한 기술적 분석 관련 연구가 활발하게 진행되었다 (Lo 등, 2000; Leigh 등, 2002; Leigh 등, 2002; Leigh 등, 2004).

기술적 분석을 이용하는 투자자들은 매매 타이밍을 결정하기 위하여 주가 차트를 분석하는 방법을 사용하며 (Deboeck, 1994), 이 방법을 패턴 분석이라고 한다. 과거 주가 차트에 나타나는 특정 모양들을 연구하여 현재 주가의 움직임을 맞추어 봄으로써 주가의 패턴을 예측하는 방법이다. 주가의 패턴 분석을 통한 주가 예측 시스템 개발을 위해 유사한 패턴을 찾는 방법으로 유클리디안 거리 (Euclidean distance) 방법을 이용하거나 인공지능 (artificial intelligence) 방법을 사용하였다 (Kim 등, 2002; Chung 등, 2004; Dong과 Zhou, 2002). 동적 패턴 인식 알고리즘을 사용하여 유사한 패턴을 찾아내서 주식 매매 전략에 적용할 수 있는 최적의 매매 타이밍에 관한 연구는 지금까지 이루어지지 않았다.

본 연구는 시스템 트레이딩을 이용하여 선물시장에서 거래 할 때 최적의 매매 타이밍을 알아보고 이에 적합한 전략이 무엇인가를 알아보는 것이 목적이다. 패턴인식 알고리즘인 동적시간왜곡 (DTW; Dynamic Time Warping) 알고리즘을 이용하여 빈도별 (10분, 30분, 60분, 일별) 유사 패턴을 찾아내고 최적의 매매 타이밍을 알아보는 것이다. 이를 위해 주식시장의 대표적인 패턴들을 알아보고, 패턴들의 기간을 빈도별로 DTW를 이용하여 유사한 패턴들을 찾는다. 인식된 유사한 패턴들에 대하여 기술적 지표들의 개별 전략을 적용한 거래 시뮬레이션을 실시하였다. 대표적인 패턴과 유사 패턴에 전략들을 적용했을 때 어떤 빈도에서 높은 수익률을 발생시키는지 살펴보았다.

이 논문은 다음과 같이 구성되어 있다. 2절에서는 주가지수선물시장, 기술적 분석, 시스템 트레이딩, DTW 알고리즘에 대해 소개하였으며, 3장에서는 DTW를 이용한 패턴인식 알고리즘 기반의 제안 모델을 자세히 설명하였다. 4장에서는 제안 모델의 실증분석 결과를 분석하였으며, 결론에서는 본 연구의 기대효과 및 향후 연구에 대해 서술하였다.

2. 선행연구

2.1. 주가지수선물시장

선물거래란 표준화된 선물계약의 매매라고 할 수 있으며, 선물계약은 수량, 규격, 품질 등이 표준화되어 있는 상품 또는 금융자산에 대하여 현재 (계약체결시점에서)에 서로 정한 가격 (선물가격)으로 미래의 특정시점에 인수·인도할 것을 약속하는 법적 구속력을 갖는 계약으로서 일정한 거래소 (선물거래소)에서 규정한 정형화된 방법으로 거래하는 것을 말한다. 주가지수선물 (Stock Index Futures)은 선물의 원리를 주가지수라는 것에 적용한 것이다 (박인찬 등 2009). 즉, 주가지수선물거래는 증권시장에서 매매되고 있는 전체 주식 또는 일부 주식의 가격수준을 나타내는 주가지수를 대상으로 장래의 일정시기에 계약 체결 시 약정한 가격으로 결제할 것을 약속하는 거래이다 (김경재, 1997).

주가지수선물이 다른 선물에 비해 특이한 점은 거래대상이 유형의 실물이 아니므로 기술적으로 인수·인도가 불가능한 무형의 주수라는 것이다. 따라서 반대매매를 통하여 만기일 전에 포지션을 청산하든지 아니면 만기일에 가서 그 차익을 현금으로 결제한다. 이러한 주가지수 선물시장은 매매대상과 결제방법 등에서 선물거래 및 다른 선물거래와 다음과 같이 구별되는 특징이 있다. 첫째, 실물이 존재하지 않기 때문에 실물인지도가 불가능하여 현금결제 방법 채택되고 있다. 둘째, 선물시장이 효율적인 시장으로서 존속할 수 있기 위해서 결제불이행 위험성을 방지할 수 있는 제도적 장치인 결제수단, 증거금, 일일정산제도를 마련하고 있다. 증거금이란 가격 하락 시 매수자의 계약위반 가능성으로부터 매도자를 보호하고 가격 상승 시 매도자의 계약위반 가능성으로부터 매수자를 보호하는 제도로서 모든 선물거래 참여자들이 계약을 성실히 이행하겠다는 신용의 표시로 선물거래 중개회사를 통하여 결제기관에 납부하는 금액이다. 일일정산제도는 매일매일 장이 종료된 후 청산소는 새로운 당일의 정산가격을 기준으로 보유하고 있는 포지션에 대한 가격변동위험을 평가하여 납부해야 할 증거금을 산정하고 이를 다시

전일의 증거금 잔액과 비교하여 그 차액을 산정한다. 셋째, 선물거래는 계약수행을 위한 수단으로서 최종결제일 이전의 계약기간 중 언제든지 거래소의 시장가격으로 반대매매가 가능하며, 반대매매 후 최종 순익에 따른 차금결제로 거래를 종결시킬 수 있어 편리하다. 넷째, 일반적으로 거래를 위해 대납하는 증거금이 규모와 총 거래금액에 비해 매우 작기 때문에 가격등락에 따른 투자수익과 투자위험이 현금시장에서보다 훨씬 높다는 점 등이다 (변중국, 1993).

2.2. 기술적 분석

기술적 분석이란 어떤 상품의 수요와 공급에 영향을 미치는 외부요인을 연구하기보다는 그 외부의 영향으로 나타나는 시장가격 그 자체를 연구함으로써, 앞으로의 가격변화를 알아내고자하는 가격 예측기법이다 (한국선물학회, 2000). 즉, 과거 주가의 움직임을 그림이나 차트로 그려 현재 상황과 비교하여 향후 주가의 변동을 예측하는 방법이다 (장재건, 1996). 이는 주가가 과거의 패턴대로 움직인다는 가정에 입각하여 분석하기 때문에 미래의 주가와 과거의 주가가 긴밀한 상관관계를 가지는 경우에는 주가에 대해 큰 도움이 되지만, 주가란 미래의 발생 가능한 돌발적인 사건에 따라 움직일 수 있으므로 기술적 분석에 의한 주가예측은 한계점이 있다.

주식, 선물 시장에서의 기술적 지표란 기본적인 시장의 가격이나 거래량 정보들을 이용하여 시장의 장단기의 추세와 같은 시계열 정보를 좀 더 알기 쉬운 형태로 나타낼 수 있도록 변형하여 사용할 수 있게 한 것이다 (Achelis, 1995). 일반적으로 기술적 지표들은 과거의 가격 흐름과 거래량 등의 시장변수를 변형하여 시장의 패턴을 파악하고 정형화하여 미래 주가의 흐름을 예측하는 데 사용되며, 주로 최적의 매매시점을 찾아내는 목적으로 활용된다. 선물 시장에서는 다양한 종류의 기술적 지표들이 사용된다. 본 연구에서는 16개의 기술적 지표를 사용하였으며, 각각의 기술적 지표에 대한 설명과 매매 전략은 Lee 등 (2010)의 연구에서 Appendix에 자세히 설명되어 있다.

2.3. 시스템 트레이딩

시스템 트레이딩이란 체계적인 거래를 위한 기계적인 거래전략을 의미하는 것으로 시장진입 규칙 (매수, 매도규칙)과 청산규칙 (이익확보, 손실제한 포함)으로 이루어져 있다. 자동매매의 의미가 시스템 트레이딩의 개념은 아니며 모든 매매 의사결정을 마치 프로그램으로 만든 것처럼 규칙화시킨 것을 의미한다. 즉 시스템 트레이딩의 정확한 의미는 어떠한 시뮬레이션을 통한 매매규칙을 갖고 특정한 과거데이터를 이용하여 검증해본 후 그 매매 규칙을 가지고 실전에 적용하여 매매하는 것을 의미한다. 이러한 시스템 트레이딩을 하는 이유는 감정적인 의사가 개입된 매매 방법으로 수익을 확보하기가 어렵고 시장의 변동성에 의한 감정의 동요로 인해 손실이 발생하기 때문이다. 일정한 규칙을 정한 매매 방법은 거래의 일관성을 유지하고 감정에 의한 뇌동매매 손실을 최소화하기 위해 사용한다.

일반적으로 매매를 실행하는데 있어서 가장 중요한 사항은 시장진입시점을 어떻게 결정할 것인가, 일단 포지션을 보유하게 되면 어느 시점에서 이익을 취하면서 포지션을 청산할 것인가, 그리고 시장의 가격흐름이 보유포지션과 반대방향으로 진행되는 경우 어느 정도의 손실을 감수하면서 포지션을 청산할 것인가를 결정하는 일이다. 시스템 트레이딩은 순수하게 기술적 분석들만으로 거래에 관련된 모든 사항 (진입, 퇴출, 재진입, 손실한도의 설정, 자금관리, 위험관리 등)을 하나의 시스템으로 작성, 이를 충분한 과거 데이터로 실험한 후에 시스템에서 제공하는 매수·매도 신호에 의해 기계적으로 거래가 이루어진다. 하지만 주관적 거래는 거래자 개인의 과거 거래경험과 기본적 분석 및 개인적인 선호를 가지고 있는 기술적 분석 등으로 개인적으로 시장을 분석하고 거래자 자신의 판단에 따라 거래가 이루어진다.

2.4. 동적시간왜곡 알고리즘

동적시간왜곡 (DTW, Dynamic Time Warping) 알고리즘은 음성인식, 생물정보공학, 온라인 필기체 문자 인식 등에서 널리 사용되고 있는 것으로, 두 시계열 간의 거리를 최소화하는 하는 방향으로 움직이면서 매칭시켜 각 템플릿과의 누적 거리를 계산하여 최소가 되는 클래스로 인식한다. 그림 2.1을 살펴보면 DTW를 이용하여 두 시계열을 매칭시켰을 경우 유클리디안 거리 방법을 사용했을 때와는 달리 부분적으로 왜곡되거나 변형된 파형에 대해서도 적절하게 매칭시킬 수 있다는 것을 보여준다 (Keogh와 Pazzani, 1999).

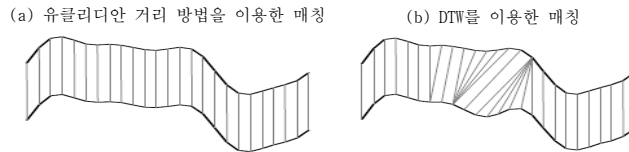


그림 2.1 두 시계열 간의 거리 매칭 (Keogh와Pazzani, 1999)

길이가 각각 m, n 인 두 개의 시계열 $X = (x_1, x_2, \dots, x_m)$ 와 $Y = (y_1, y_2, \dots, y_n)$ 가 주어졌을 때, 두 시계열의 정렬을 위해 $m \times n$ 행렬이 만들어진다 (그림 2.2). 이 행렬의 (i, j) 번째 요소는 두 점 x_i 와 y_j 간의 거리 $d(x_i, y_j)$ 를 포함한다. 두 점 x_i 와 y_j 간의 거리 $d(x_i, y_j)$ 는 유클리디안 거리 방법 $(x_i - y_j)^2$ 을 사용하여 계산한다.

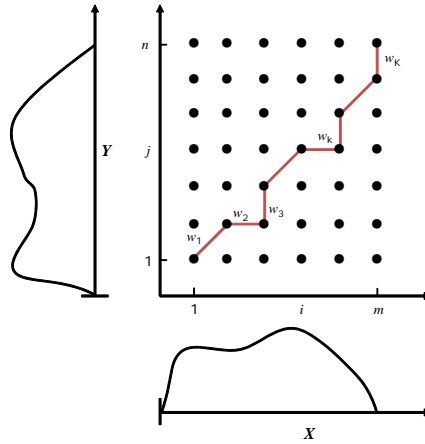


그림 2.2 두 시계열의 행렬관계와 warping 경로의 예 (Keogh와Pazzani, 1999)

warping 경로 W 는 X 와 Y 사이의 매핑을 정의하는데, W 의 k 번째 요소는 $w_k = (i, j)_k$ 로 정의된다. 따라서 $W = w_1, w_2, \dots, w_k, \dots, w_K$ 일 때, $\max(m, n) \leq K < m + n - 1$ 이다. 이 때, W 는 경계조건 (boundary condition), 연속성 (continuity), 단조성 (monotonicity)의 세 가지 조건 (Keogh와 Pazzani, 1999)을 만족해야 한다. 이 조건을 만족하는 여러 가지 warping 경로들 중에서 각각의 warping 거리

w_k 들의 총합이 최소가 되는 warping 경로를 발견해야 한다.

$$DTW(X, Y) = \frac{1}{K} \sqrt{\sum_{k=1}^K w_k} \tag{2.1}$$

여기서 K 는 서로 다른 길이를 가지고 있는 warping 경로들에 대한 보상을 위해 사용된다.

$i = j = 0$ 에서부터 시작하여 k 번째 warping 경로 w_k 의 누적 warping 거리 $D(i, j)$ 는 아래 식 (1.2)로 정의할 수 있다.

$$D(i, j) = d(x_i, y_j) + \min \left\{ \begin{array}{l} D(i-1, j-1) \\ D(i-1, j) \\ D(i, j-1) \end{array} \right\} \tag{2.2}$$

누적거리 $D(i, j)$ 는 $i = j = 0$ 에서부터 시작하며, 최종적인 유사도를 나타내는 값으로 DTW에서 최단 경로를 결정하게 되는 측정값이 된다.

3. 제안모델

본 절에서는 주식거래 시 주가데이터의 빈도에 따른 최적의 매매 타이밍을 알아보기 위한 본 연구의 제안 모델에 대해 설명한다 (그림 3.1).

제안 모델의 1단계에서는 Lee 등 (2010)의 연구에서 제시된 주식시장에 존재하는 대표적인 6개의 패턴을 확인한다. 6개의 대표 패턴은 단기상승, 단기하락, 대세상승, 대세하락, 천정횡보, 바닥횡보 등이며 자세한 설명은 표 3.1에 제시하였다. 6개의 대표 패턴마다 빈도별 주가 데이터 (10분, 30분, 60분, 일별 데이터)를 생성한다. 예를 들어, 단기상승 패턴의 기간을 10분, 30분, 60분, 일별로 데이터로 구성된 데이터 세트를 생성한다. 따라서 6개의 대표 패턴에 각각 4개의 빈도별 데이터 24개가 생성된다. 2단계에서는 6개의 패턴이 가지고 있는 빈도별 데이터와 유사한 패턴을 과거 주식시장에서 각각 찾는다. 이때 유사패턴을 찾기 위해 2절에서 설명된 DTW 알고리즘이 사용되며, 6개의 패턴 기간과 중복되지 않는 유사 기간을 데이터 빈도별로 각각 찾아낸다. 마지막 단계에서는 6개의 대표 패턴 기간과 DTW가 인식한 유사 패턴 기간에서 거래 시뮬레이션을 실시한다. 6개의 대표 패턴 기간과 유사 패턴 기간에 기술적 지표들의 개별 전략을 적용하고 빈도별 수익률을 비교한다.

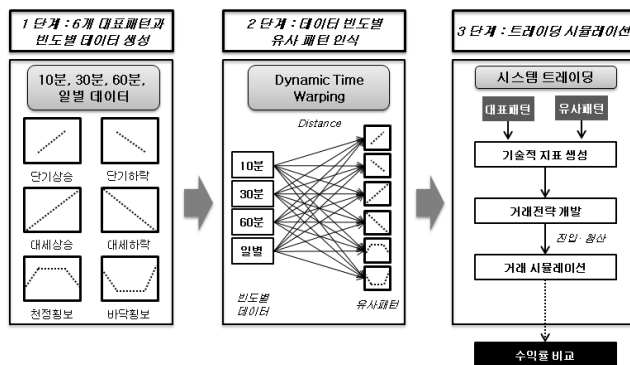


그림 3.1 본 연구의 제안 모델

표 3.1 본 연구에서 사용된 6개 대표 패턴 (Lee 등, 2010)

패턴	설명
단기상승	약 2~3주 동안 주가의 지속적 상승
단기하락	약 2~3주 동안 주가의 지속적 하락
대세상승	약 6개월 동안 주가의 지속적 상승
대세하락	약 6개월 동안 주가의 지속적 하락
천정 횡보	상승 추세 이후 횡보하는 움직임
바닥 횡보	하락 추세 이후 횡보하는 움직임

4. 실증분석

한국의 선물시장은 1996년 5월 3일 개장 이래 다양한 요인들의 영향으로 역동적인 모습을 나타내고 있다 (변현우 등, 2009). 선물시장의 예측 불가능한 변동성으로 인해 트레이더 같은 전문 투자자들은 인간으로써 한정된 능력을 가지고 있는 자신들에게 기술적으로 지원해 줄 수 있는 도구가 필요함을 느끼게 되었다 (Lee 등, 2010). 따라서 본 연구에서는 선물거래 시 기술적 분석에 주로 사용되는 패턴 분석을 수행 할 때 어떠한 빈도의 데이터가 유용한지를 분석하고자 한다. 실험에 사용된 데이터는 1996년 5월 3일부터 2010년 12월 30일까지 KOSPI 200 연결선물지수 (10분, 30분, 60분, 일별)이며, 그림 4.1은 과거 15년 (전체 실험기간) 동안의 KOSPI 200의 흐름을 보여주고 있다.

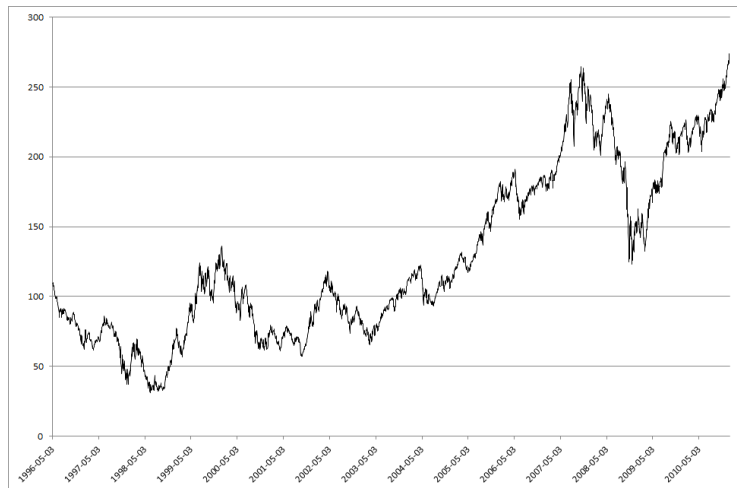


그림 4.1 1996년 5월 3일부터 2010년 12월 30일까지 KOSPI 200 선물지수의 전체 흐름

전체 실험 기간에서 훈련기간은 1996년 5월 3일부터 2000년 2월 3일까지이다. 훈련기간에서 Lee 등 (2010)이 제시한 우리나라 선물시장의 대표 패턴으로 단기상승, 단기하락, 대세상승, 대세하락, 천정횡보, 바닥횡보 등의 기간을 선택하였으며, 각각의 기간은 표 4.1에 설명되어 있고 각 패턴들의 형태는 그림 4.2에서 보여주고 있다. 본 연구에서 사용된 6개의 대표패턴 기간은 실거래일 (영업일)수를 기준으로 816일이며 전체 기간 3,533일의 약 23%의 비중을 차지하고 있다. 대표 패턴기간과 중복되지 않는 유사패턴의 기간을 고려하면 6개의 대표 패턴은 총 1,632일로써 전체기간의 약 46%이다. 이것은 우리나라 선물시장의 전체기간 중 절반이 패턴을 가지고 움직인다는 것을 말하며, 6개의 대표 패턴을 이용한 분석이 의미를 가지고 있다는 것을 알 수 있다. 검증 기간은 2000년 2월 7일부터 2010년 12월 30일

이며 이 구간에서 6개 패턴들과 유사한 패턴들을 빈도별로 찾은 후 대표 패턴과 유사 패턴에 대한 거래 시뮬레이션이 수행된다. 거래 시뮬레이션에 사용된 수익률은 보유한 자산이 일 년간 창출한 연간 수익률을 의미하며, 실제 거래 상황을 고려하여 거래 시뮬레이션에 수수료 10,000원과 슬리피지 (slippage) 25,000원을 적용하였다. 수수료는 위탁중개회사가 위탁자로부터 주문접수 및 체결에 대한 용역에 대한 대가로 받는 금액이며, 슬리피지란 시장가주문을 넣었을 때 주문을 넣은 시점의 가격으로 체결되지 않고 더 나쁜 가격으로 체결되는 것이다.

표 4.1 6개 대표 패턴들의 기간 (Lee 등, 2010)

패턴	시작일	마지막일
단기상승	1997년 12월 26일	1998년 3월 5일
단기하락	1998년 3월 6일	1998년 6월 15일
대세상승	1998년 10월 1일	1999년 7월 9일
대세하락	1996년 9월 9일	1997년 12월 24일
전정 횡보	1999년 7월 12일	2000년 2월 3일
바닥 횡보	1998년 6월 16일	1998년 9월 30일

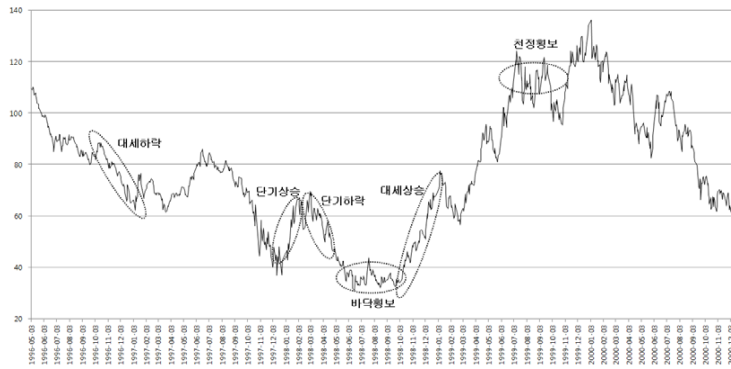


그림 4.2 6개의 대표 패턴들의 형태

Lee 등 (2010)이 제안한 6개의 패턴들과 유사한 패턴들을 찾기 위해, 2절에서 설명한 DTW 알고리즘을 사용하여 빈도별 패턴인식을 실시하였으며, DTW를 이용한 거리 계산 (식 1.2)을 위해 Visual Basic Programming을 사용하였다. KOSPI 200 데이터는 연속형이므로 패턴인식을 위해 데이터의 표준화가 필요하며, 아래 식 (4.1)에 의해 차분된 데이터를 사용하여 유사패턴과의 거리를 계산했다.

$$\Delta X = X_t - X_{t-1} \tag{4.1}$$

X_t 는 일정 시점의 KOSPI 200 가격이고 X_{t-1} 는 일정 시점 이전의 KOSPI 200 가격이다.

표 4.2는 DTW를 이용하여 찾은 데이터 빈도별 (10분, 30분, 60분, 일별) 6개의 대표 패턴들과 유사한 패턴들의 기간과 거리 (유사도)를 보여주고 있다. 동일한 패턴 기간이지만 데이터 빈도별로 인식한 유사패턴의 기간이 다르다는 것을 알 수 있다. 즉, 단기상승 패턴에서 10분 데이터의 경우 2001년 6월 15일에서 2001년 8월 14일 (거리 162.90), 30분 데이터의 경우 2001년 9월 17일에서 2001년 11월 16일 (거리 100.40), 60분 데이터의 경우 2003년 7월 21일에서 2003년 9월 22일 (거리 76.60), 일별 데이터의 경우 2001년 9월 27일에서 2001년 11월 28일 (거리 33.35)이 유사 패턴으로 인식되었다. 데이터 빈도

별로 단기상승 패턴을 구분했을 때 다른 유사 패턴 기간이 매칭 되는 것을 보여준다. 나머지 패턴들 (단기하락, 대세상승, 대세하락, 천정횡보, 바닥횡보)도 단기상승 패턴과 마찬가지로 데이터 빈도별로 조금씩 다른 유사 패턴 기간들과 매칭 되는 것을 알 수 있다.

표 4.2 DTW를 이용하여 찾은 데이터 빈도별 유사패턴

단기상승			
데이터 빈도	시작날짜	마지막날짜	거리
10분	2001년 6월 15일	2001년 8월 14일	162.90
30분	2001년 9월 17일	2001년 11월 16일	100.40
60분	2003년 7월 21일	2003년 9월 22일	76.60
일별	2001년 9월 27일	2001년 11월 28일	33.35
단기하락			
데이터 빈도	시작날짜	마지막날짜	거리
10분	2001년 5월 31일	2001년 9월 7일	230.15
30분	2001년 6월 4일	2001년 9월 11일	149.50
60분	2001년 6월 11일	2001년 9월 17일	107.30
일별	2001년 5월 23일	2001년 8월 30일	44.90
대세상승			
데이터 빈도	시작날짜	마지막날짜	거리
10분	2001년 2월 16일	2001년 11월 23일	894.18
30분	2001년 2월 22일	2001년 11월 26일	523.20
60분	2001년 2월 20일	2001년 11월 22일	391.64
일별	2003년 7월 3일	2004년 4월 12일	177.00
대세하락			
데이터 빈도	시작날짜	마지막날짜	거리
10분	2001년 1월 16일	2003년 5월 3일>	810.52
30분	2002년 6월 20일	2003년 10월 6일	653.25
60분	2004년 6월 1일	2005년 10월 6일	545.85
일별	2003년 1월 28일	2004년 5월 14일	253.20
천정횡보			
데이터 빈도	시작날짜	마지막날짜	거리
10분	2001년 4월 24일	2001년 11월 20일	539.20
30분	2001년 4월 30일	2001년 11월 23일	376.00
60분	2001년 5월 03일	2001년 11월 27일	283.14
일별	2004년 12월 16일	2005년 8월 12일	132.95
바닥횡보			
데이터 빈도	시작날짜	마지막날짜	거리
10분	2003년 4월 29일	2003년 8월 20일	214.30
30분	2003년 5월 12일	2003년 8월 29일	140.65
60분	2001년 5월 24일	2001년 9월 11일	106.65
일별	2001년 6월 28일	2001년 10월 18일	43.80

빈도별 데이터의 유용성을 알아보기 위해 각각의 대표 패턴들 기간과 유사패턴들 기간에서 빈도별 데이터를 이용한 거래 시뮬레이션을 실시하였으며, 이때 이용된 소프트웨어는 *Tradestation 2000i*이다. *Tradestation 2000i*는 미국의 Omega Research사가 개발하였으며, 언제 매수하고 언제 매도할 것인가에 대한 분석 기능을 지원하는 시스템 트레이딩 툴이다. 이러한 테스트를 기초로 만들어진 투자자만의 투자 전략에 맞는 매매 시점이 도래하면 차트 창에 경보 기능을 주어 쉽게 매매 타이밍을 인식할 수 있도록 지원하는 소프트웨어이다. 표 4.3은 대표 패턴 기간에서의 수익률 상위 5위의 지표들과 수익률의 빈도별 (10분, 30분, 60분, 일별) 수익률이다. 거래 시뮬레이션에는 16개의 기술적 지표들의 개별 전략 (Lee 등, 2010)이 사용되었다. 단기상승 패턴에서 일별 데이터의 경우 거래가 발생하지 않았다. 표 4.4는 6개의 대표 패턴들과 유사한 패턴들의 빈도별 거래 시뮬레이션 결과를 보여주고 있다. 대표 패턴 기간과 유사 패턴 기간에서 빈도별로 수익률 상위를 보이는 기술적 지표들의 전략이 다르다는 것을 볼 수 있으며, 수익률 또한 다르다는 것을 알 수 있다. 이것은 데이터 빈도별로 거래 타이밍 즉, 매수, 매도 기회가 다르기 때문이라는 것을 의미한다.

표 4.3 6개 대표 패턴 기간에서 수익률 상위 5위의 기술적 지표와 빈도별 수익률(%)

단기상승								
순위	10분		30분		60분		일별	
	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률
1	CCI	86.60	Stochastic	23.45	Band%b	36.95	Band%b	0.00
2	CO	33.55	RSI	16.75	Stochastic	31.95	CCI	0.00
3	SMI	7.00	PO	14.05	NCO	27.15	CO	0.00
4	DMI	-5.45	CO	9.20	Momentum	25.15	DMI	0.00
5	Williams%R	-7.30	TRIX	-10.25	SMI	24.85	ROC	0.00

단기하락								
순위	10분		30분		60분		일별	
	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률
1	Band%b	61.35	CO	40.20	CO	60.25	SMI	13.45
2	CCI	16.95	Momentum	38.55	SMI	38.05	RSI	1.05
3	CO	1.00	NCO	35.40	MACD	29.60	NCO	-0.35
4	DMI	-12.50	MACD	9.60	RSI	28.10	PO	-0.35
5	ROC	-15.95	SMI	6.35	Momentum	27.15	Stochastic	-0.35

대세상승								
순위	10분		30분		60분		일별	
	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률
1	TRIX	237.65	SMI	213.55	DMI	104.10	Momentum	53.60
2	MACD	108.70	Momentum	168.20	CCI	101.25	PO	-2.10
3	PO	104.75	CCI	136.00	CO	101.05	MACD	-2.45
4	CO	101.35	MACD	109.00	Williams%R	66.10	TRIX	-2.45
5	CCI	0.65	TRIX	70.55	RSI	13.00	VO	-2.45

대세하락								
순위	10분		30분		60분		일별	
	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률
1	CO	119.65	Momentum	91.45	CCI	67.60	RSI	14.80
2	CCI	84.45	SMI	86.70	Stochastic	61.25	Momentum	7.45
3	SMI	45.90	MACD	66.65	VO	58.25	Stochastic	-0.35
4	DMI	32.15	CCI	28.75	PO	40.55	CO	-0.70
5	Williams%R	-0.65	DMI	28.40	MACD	33.05	PO	-0.70

천정정보								
순위	10분		30분		60분		일별	
	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률
1	DMI	156.70	SMI	207.05	CCI	222.75	SMI	28.15
2	TRIX	24.30	Momentum	170.75	NCO	158.55	RSI	6.80
3	Stochastic	23.10	NCO	155.80	ROC	139.05	CCI	6.30
4	CO	10.55	CCI	105.85	CO	97.55	Stochastic	-1.05
5	CCI	-2.65	CO	88.05	Band%b	74.60	PO	-2.10

바닥횡보								
순위	10분		30분		60분		일별	
	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률
1	CO	21.35	SMI	36.30	PO	32.30	RSI	1.40
2	TRIX	21.00	CCI	35.25	MACD	6.15	VO	0.70
3	MACD	8.05	Stochastic	3.25	CO	4.35	Stochastic	-0.35
4	PO	3.80	TRIX	-4.40	RSI	2.00	CO	-0.70
5	SMI	-0.20	MACD	-8.90	VO	1.15	DMI	-0.70

표 4.5는 6개 패턴 기간에서 수익률이 가장 높은 5개 지표들의 데이터 빈도별 평균 수익률을 보여주 고 있다. 단기상승, 단기하락 패턴에서는 60분 데이터의 수익률이 가장 높았으며, 대세상승, 대세하락, 천정정보, 바닥횡보 패턴에서는 30분 데이터의 수익률이 가장 높다. 하지만 60분 데이터의 수익률과 큰 차이를 보이지 않았다. 10분 데이터의 경우 거래세로 인해 수익을 저해하는 것을 알 수 있었으며, 일별 데이터의 경우 Lee 등 (2010)이 언급했던 것처럼 갭 상승이나 갭 하락으로 인한 손실이 나타나는 것을 알 수 있었다. 주가 갭이란 어떤 주식이 특정한 전일의 고가보다 높이 시작하여 그대로 계속 상승하여 양봉을 유지하여 전일과 당일 일봉 사이에 공간이 발생하는 것을 갭 상승이라고 한다. 반대로 전일의 저가보다 낮은 가격으로 떨어진 후 계속 하락하는 경우가 발생하는 것을 갭 하락이라고 한다. 표 4.5에서 의 평균은 빈도별 6개 대표 패턴들의 수익률을 평균한 것을 말하는데, 30분 데이터가 65.72%로 가장 높

표 4.4 유사 패턴기간에서 수익률 상위 5위의 기술적 지표와 빈도별 수익률 (%)

단기상승								
순위	10분		30분		60분		일별	
	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률
1	Stochastic	19.30	CO	30.40	CO	17.95	SMI	65.08
2	DMI	10.55	DMI	8.90	CCI	10.85	CO	38.50
3	MACD	4.15	Stochastic	8.60	SMI	4.80	CCI	14.50
4	SMI	3.10	RSI	7.20	Stochastic	2.60	DMI	-77.50
5	TRIX	2.65	PO	3.50	RSI	1.80	Williams%R	-79.38

단기하락								
순위	10분		30분		60분		일별	
	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률
1	PO	34.15	DMI	28.55	RSI	40.00	SMI	7.55
2	SMI	27.00	PO	25.80	DMI	15.00	RSI	2.40
3	DMI	18.65	RSI	19.85	NCO	12.10	PO	-0.70
4	MACD	3.35	CO	8.75	VO	10.95	CO	-1.05
5	TRIX	1.10	Stochastic	1.10	SMI	3.80	VO	-1.05

대세상승								
순위	10분		30분		60분		일별	
	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률
1	PO	141.39	TRIX	50.40	TRIX	31.85	Momentum	34.00
2	RSI	93.80	PO	46.55	PO	24.85	RSI	-1.00
3	DMI	59.14	RSI	29.75	DMI	17.15	Stochastic	-2.45
4	Stochastic	52.85	Stochastic	29.40	RSI	16.44	PO	-3.15
5	Williams%R	-142.10	DMI	24.85	Stochastic	15.40	TRIX	-3.85

대세하락								
순위	10분		30분		60분		일별	
	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률
1	DMI	44.00	CO	73.70	PO	51.45	Momentum	20.05
2	Band%b	43.50	DMI	39.30	DMI	27.55	TRIX	7.00
3	CO	30.25	PO	29.85	SMI	26.55	PO	5.60
4	CCI	8.30	CCI	22.85	CO	21.00	Stochastic	5.25
5	ROC	-76.80	Stochastic	9.65	MACD	20.75	VO	-7.35

전경정보								
순위	10분		30분		60분		일별	
	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률
1	RSI	70.00	Stochastic	38.55	PO	54.10	CO	2.80
2	DMI	43.04	RSI	32.30	RSI	7.90	TRIX	2.80
3	Stochastic	38.50	PO	15.10	DMI	5.40	VO	2.80
4	PO	-103.59	DMI	11.05	SMI	3.25	VROC	2.80
5	Williams%R	-106.40	CO	9.50	CO	-6.20	PO	1.75

바닥정보								
순위	10분		30분		60분		일별	
	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률	기술적 지표	수익률
1	TRIX	54.25	TRIX	16.80	NCO	15.10	DMI	0.00
2	RSI	37.10	PO	16.45	DMI	5.90	CCI	-0.70
3	DMI	25.55	RSI	16.10	Stochastic	4.45	MACD	-0.70
4	Stochastic	23.80	Stochastic	12.94	RSI	4.05	NCO	-0.70
5	PO	-55.30	DMI	10.85	SMI	1.25	PO	-0.70

고 60분 데이터가 57.13%로 두 번째 높은 수익률을 보여주고 있다. 표 4.6은 유사 패턴들에 대한 빈도별 수익률 상위 5개 지표들의 평균 수익률을 보여주고 있다. 대표 패턴들과 마찬가지로 30분 데이터의 수익률이 가장 높은 것을 볼 수 있다. 즉, 선물거래를 위해 사용될 수 있는 가장 유용한 데이터 빈도는 30분 데이터이다. 대표패턴 기간에서 수익률이 높은 지표들의 전략이 유사패턴 기간에서도 높은 수익률을 가져온다는 것은 동적으로 움직이는 주가의 패턴인식을 위해 DTW 알고리즘이 유용하다는 것을 의미한다.

5. 결론

주가지수 선물시장은 1996년에 개설된 이래 약 15년 동안 세계적인 금융시장으로 지속적인 성장을

표 4.5 대표 패턴들에 대한 빈도별 수익률 상위 5개 지표들의 평균 수익률(%)

대표 패턴	10분	30분	60분	일별
단기상승	22.88	10.64	29.21	0.00
단기하락	10.17	26.02	36.63	2.69
대세상승	110.62	139.46	77.10	8.83
대세하락	56.30	60.39	52.14	4.10
천정횡보	42.40	145.50	138.50	7.62
바닥횡보	10.80	12.30	9.19	0.07
평균	42.20	65.72	57.13	3.89

표 4.6 유사 패턴들에 대한 빈도별 수익률 상위 5개 지표들의 평균 수익률(%)

대표 패턴	10분	30분	60분	일별
단기상승	7.95	16.81	7.60	-7.76
단기하락	16.85	36.19	16.37	1.43
대세상승	41.02	35.07	21.14	4.71
대세하락	9.85	29.46	22.43	6.11
천정횡보	-11.69	12.89	11.54	2.59
바닥횡보	17.08	6.15	6.15	-0.56
평균	13.51	22.76	15.60	1.09

해왔다. 하지만 이 기간 동안 수많은 변동성이 존재하였고, 시장 참여자들은 변동성을 분석하기 위해 다양한 기술적 분석을 개발하고 적용하였다.

본 연구에서는 주식거래 시 고려되는 주가 데이터의 빈도에 따른 매매 기회 분석을 위해 주식시장을 대표하는 6개의 패턴들을 대상으로 빈도별 거래 시뮬레이션을 실시하였다. 그 결과 대부분의 패턴에서 30분 데이터를 사용했을 때 높은 수익률을 얻을 수 있다는 것을 알 수 있었다. 또한 DTW를 이용하여 유사한 패턴을 찾고 동일한 거래 시뮬레이션을 수행한 결과 30분 데이터의 수익률이 가장 우수한 것을 볼 수 있었다. 실험 결과는 시장 참여자들이 주식시장을 분석하거나 거래전략을 수립하기 위해 수행하는 백테스팅을 실시할 때나 실제 거래 시 데이터의 빈도 선정에 있어 유용한 정보를 제공해 줄 것이다.

본 연구는 동적 패턴 인식 알고리즘을 사용하여 주가지수선물시장에서의 최적매매기회를 탐색하고 거래 전략을 개발하였다. 하지만 전 세계적으로 다양한 패턴을 가진 금융시장이 존재하고 있다. 따라서 본 연구에서 제시한 방법론이 한국의 금융시장 뿐 아니라 다른 금융시장에도 적용 가능한지 고려되어야 할 것이다.

참고문헌

- 박인찬, 권오진, 김태윤 (2009). 시계열 모형을 이용한 주가지수 방향성 예측. <한국데이터정보과학회지>, **20**, 991-998.
- 김경재 (1997). <범주형 전처리과정을 이용한 인공지능 기법에 의한 주가지수선물가격의 예측: KOSPI200 선물 시장을 중심으로>, 석사학위논문, 한국과학기술원, 서울.
- 변종국 (1993). 주가지수선물의 성격과 이해. <증권>, **78**, 26-48.
- 변현우, 송치우, 한성권, 이태규, 오경주 (2009). 변동성 지수기반 유전자 알고리즘을 활용한 계층구조 포트폴리오 최적화에 관한 연구. <한국데이터정보과학회지>, **20**, 467-478.
- 신양규 (2009). 글로벌경제위기에서 콜금리와 환율의 인과관계에 관한 연구. <한국데이터정보과학회지>, **20**, 655-660.
- 장재건 (1996). <기술적 분석지표를 이용한 선물투자기법>, 진리탐구, 서울.
- 한국선물학회 (2000). <선물 이론 및 실무>, 이론편, 금빛서원, 서울.
- Achelis S. B. (1995). *Technical analysis from A to Z*, Probus Publishing, Chicago.

- Aziz, A. M., Tummala, M. and Cristi, R. (1999). Fuzzy logic data correlation approach in multisensor-multitarget tracking systems. *Signal Processing*, **76**, 195-209.
- Castellanos, J. A. and Tards, J. D. (1999). *Mobile robot localization and map building: A multisensor fusion approach*, Kluwer Academic Publishers, Boston.
- Chung, F. L., Fu, T. C., Ng, V. and Robert, W. P. (2004). An evolutionary approach to pattern-based time series segmentation. *IEEE Transactions on Evolutionary Computation*, **8**, 471-489.
- Deboeck, G. (1994). *Trading on the edge*, John Wiley & Sons, Inc, Canada.
- Dong, M. and Zhou, X. S. (2002). Exploring the fuzzy nature of technical patterns of U.S stock market. *ICONIP'02-SEAL'02-FSKD'02*, Singapore, 18-22.
- Gil, J. W. (2003). The Return generating process of corporate bonds based on credit ratings. *Journal of Korean Data & Information*, **14**, 805-815.
- Jouseau, E. and Dorizzi, B. (1999). Neural network and fuzzy data fusion: Application to an on-line and real-time vehicle detection system. *Pattern Recognition Letters*, **20**, 97-107.
- Keogh, E. J. and Pazzani, M. J. (1999). Scaling up dynamic time warping to massive dataset. '99 *Proceedings of the Third European Conference on Principles of Data Mining and Knowledge Discovery*.
- Keogh, E. J. and Pazzani, M. J. (2001). Derivative dynamic time warping. *First SIAM International Conference on Data Mining*, Chicago, USA.
- Kim, S. D., Lee, J. W., Lee, J. W. and Chae, J. S (2002). A two-phase stock trading system using distributional differences. *Lecture Notes in Computer Science*, **2453**, 399-423.
- Lee, S. J., Ahn, J. J., Oh, K. J. and Kim, T. Y. (2010). Using rough set to support investment strategies of real-time trading in futures market. *Applied Intelligence*, **32**, 364-377.
- Leigh, W., Modani, N. and Hightower, R. (2004). A computational implementation of stock charting: A abrupt volume increase as signal for movement in New York stock exchange composite index. *Decision Support Systems*, **37**, 515-530.
- Leigh, W., Modani, N., Purvis, R. and Roberts, T. (2002). Stock market trading rule discovery using technical charting heuristics. *Expert Systems with Applications*, **23**, 155-159.
- Leigh, W., Purvis, R. and Ragusa, J. M. (2002). Forecasting the NYSE composite index with technical analysis, pattern recognizer, neural network, and genetic algorithm: A case study in romantic decision support. *Decision Support Systems*, **32**, 161-174.
- Lo, A. W., Mamaysky, H. M. and Wang, J. (2000). Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation. *Journal of Finance*, **55**, 1705-1770.
- Morriss, S. B. (1994). *Automated manufacturing systems - Actuators, controls, sensors and robotics*, McGraw-Hill, Glencoe.
- Opitz, F., Henrich, W. and Kausch, T. (2004). Data fusion development concepts within complex surveillance systems, *The 7th International Conference on Information Fusion*.

Finding the optimal frequency for trade and development of system trading strategies in futures market using dynamic time warping

Suk Jun Lee¹ · Kyong Joo Oh²

^{1,2}Department of Information and Industrial Engineering, Yonsei University

Received 22 January 2011, revised 6 March 2011, accepted 17 March 2011

Abstract

The aim of this study is to utilize system trading for making investment decisions and use technical analysis and Dynamic Time Warping (DTW) to determine similar patterns in the frequency of stock data and ascertain the optimal timing for trade. The study will examine some of the most common patterns in the futures market and use DTW in terms of their frequency (10, 30, 60 minutes, and daily) to discover similar patterns. The recognized similar patterns were verified by executing trade simulation after applying specific strategies to the technical indicators. The most profitable strategies among the set of strategies applied to common patterns were again applied to the similar patterns and the results from DTW pattern recognition were examined. The outcome produced useful information on determining the optimal timing for trade by using DTW pattern recognition through system trading, and by applying distinct strategies depending on data frequency.

Keywords: Dynamic time warping, futures market, technical analysis, technical indicator.

¹ Ph.D. Candidate, Department of Information and Industrial Engineering, Yonsei University, Seoul 120-749, Korea.

² Corresponding Author : Associate Professor, Department of Information and Industrial Engineering, Yonsei University, Seoul 120-749, Korea. E-mail: johanoh@yonsei.ac.kr