



## 저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

박사학위 논문

주가 패턴과 유전자 알고리즘을 이용한 주식  
매매 시스템 개발

Development of a Stock Trading System Using  
Stock Price Patterns and Genetic Algorithms



국민대학교 비즈니스IT전문대학원  
트레이딩시스템전공

양 훈 석

2018

주가 패턴과 유전자 알고리즘을 이용한

주식 매매 시스템 개발

Development of a Stock Trading System Using  
Stock Price Patterns and Genetic Algorithms

지도교수 김 선 응

이 논문을 경영정보학 박사학위 청구논문으로 제출함

2019 년 6 월

국민대학교 비즈니스IT전문대학원

트레이딩시스템전공

양 훈 석

2018

梁 薰 錫 의

經營情報學博士學位 請求論文을 認准함

2019년 6월

審査委員長 최 흥 식



審査委員 김 선 웅



審査委員 안 현 철



審査委員 김 인 재



審査委員 김 충 영



國民大學校

비즈니스 IT 專門大學院

# 차 례

차례	i
그림 차례	ii
표 차례	iii
국문 요약	iv
제1장 서론	1
제2장 이론적 배경	4
2.1 추가예측을 위한 선행 연구들	4
2.2 패턴 분석	6
2.3 M&W 파동 패턴	10
2.4 유전자 알고리즘	16
2.5 전진 분석	19
제3장 시스템 개발	21
3.1 매매전략 수립	22
3.2 유전자 알고리즘 최적화	39
3.3 전진 분석 적용	42
3.4 테스트 대상 종목군 선정	44
3.5 시뮬레이션 환경 설정	46
3.6 거래내역 통합 및 재가공	46
제4장 실험 결과	52
제5장 결론 및 향후 연구	60
참고 문헌	63
영문 초록	67



# 그 립 차 례

<그림 1> 차트 패턴 . . . . .	6
<그림 2> Merrill의 M자 파동 패턴 . . . . .	11
<그림 3> Merrill의 W자 파동 패턴 . . . . .	12
<그림 4> 이중 천장, 이중 바닥의 M&W 패턴 . . . . .	16
<그림 5> 유전자 알고리즘 처리 순서도 . . . . .	17
<그림 6> 전진분석 방법 . . . . .	19
<그림 7> 시뮬레이터 구성도 . . . . .	21
<그림 8> 전환점 계산 방식의 종류 . . . . .	24
<그림 9> 지그재그 파동의 실제 적용 예 . . . . .	26
<그림 10> 4점 패턴 . . . . .	28
<그림 11> M자 패턴의 재분류 . . . . .	31
<그림 12> W자 패턴의 재분류 . . . . .	32
<그림 13> 매매전략 기본 로직 . . . . .	35
<그림 14> 시장추세를 반영한 패턴 매매 . . . . .	36
<그림 15> M&W 파동 패턴을 이용한 매매 . . . . .	37
<그림 16> 초기집단 설정 결과 예시 . . . . .	39
<그림 17> 적응도 평가 과정 예시 . . . . .	40
<그림 18> 다점 교차 방식 . . . . .	41
<그림 19> 10진수 변환 예시 . . . . .	41
<그림 20> OOS 누적수익률 . . . . .	43
<그림 21> 수익률 곡선 비교 . . . . .	57
<그림 22> 보유 종목수 변동 추이 . . . . .	58
<그림 23> 하모닉 패턴 중 가틀리, 나비, 박쥐 패턴 . . . . .	62

## 표 차 례

<표 1> 사용 변수 목록 . . . . .	38
<표 2> WFA 테스트 케이스 . . . . .	42
<표 3> 거래내역 . . . . .	43
<표 4> 고정 금액 방식으로 가공한 매매내역 . . . . .	47
<표 5> 누적 금액 방식의 일별 잔고 수량 테이블 . . . . .	49
<표 6> 종목별 진입가격 테이블 . . . . .	50
<표 7> 누적 금액 방식으로 채투자했을 때의 일별 평가금액 . . . . .	51
<표 8> 과거 테스트 수익률 비교 . . . . .	52
<표 9> 과거 테스트 수익률의 평균 및 표준편차 . . . . .	53
<표 10> 시뮬레이션 결과 . . . . .	54
<표 11> 필터링된 최적화 결과 . . . . .	56



## 국문 요약

### 주가 패턴과 유전자 알고리즘을 이용한 주식 매매 시스템 개발

국민대학교 비즈니스IT전문대학원

트레이딩시스템전공

양 훈 석

투자자들은 기업의 내재가치 분석, 기술적 보조지표 분석 등 복잡한 분석보다 차트(chart)에 나타난 그래프(graph)의 모양으로 매매 시점을 찾는 직관적인 방법을 더 선호하는 편이다. 하지만 패턴(pattern) 분석 기법은 IT 구현의 난이도 때문에 사용자들의 요구에 비해 전산화가 덜 된 분야로 여겨진다. 최근에는 인공지능(artificial intelligence, AI) 분야에서 신경망을 비롯한 다양한 기계학습(machine learning) 기법을 사용하여 주가의 패턴을 연구하는 사례가 많아졌다. 특히 IT 기술의 발전으로 방대한 차트 데이터를 분석하여 주가 예측력이 높은 패턴을 발굴하는 것이 예전보다 쉬워졌다. 지금까지의 성과로 볼 때 가격의 단기 예측력은 높아졌지만, 장기 예측력은 한계가 있어서 장기 투자보다 단타 매매에서 활용되는 수준이다. 이외에 과거 기술력으로 인식하지 못했던 패턴을 기계적으로 정확하게 찾아내는 데 초점을 맞춘 연구도 있지만 찾아진 패턴이 매매에 적합한지 아닌지는 별개의 문제이기 때문에 실용적인 부분에서 취약할 수 있다. 일레로 깃발(flags) 패턴이 나오면 이후에 가격이 상승한다고 알려졌



으며, 기존의 대다수 연구는 이렇게 가격 예측력이 있는 패턴을 정확하게 인지하는 방법을 찾는 게 목적이었다.

본 연구는 주가 예측력이 있는 패턴을 찾으려는 기존 연구 방법과 달리 발생할 수 있는 패턴들을 정의해 놓고 과거에 수익이 높았던 패턴과 같은 패턴이 나올 때 매매하는 방법을 제안한다. 패턴을 분류하는 것 자체가 대단히 복잡한 일이 될 수 있으나 이미 Merrill(1980)에 의해 5개의 꼭짓점을 이용하여 32가지로 분류한 M&W 파동 패턴 이론이 있다. 본 연구에서는 프로그래밍하기 쉽도록 패턴에 번호를 재부여하여 사용하였다. 32개 패턴 중 과거에 수익이 높았던 패턴을 찾아 조합하려면 경우의 수가 너무 많아 전수 테스트가 불가능한 수준이다. 그래서 최적화 문제와 관련한 연구들에서 가장 많이 사용되고 있는 인공지능 알고리즘(algorithm) 중 하나인 유전자 알고리즘(genetic algorithm, GA)을 이용하였다. 그리고 미래의 주가가 과거를 반영한다 해도 관계 움직이지 않기 때문에 전진 분석(walk-forward analysis, WFA)방법을 적용하여 과최적화(overfitting)의 실수를 줄이도록 하였다. 20종목씩 6개의 포트폴리오(portfolio)를 구성하여 테스트하였는데 5개의 전환점만으로도 효과적인 매매가 가능하였으며 결과적으로 패턴을 정교하게 찾지 않아도 됨을 확인하였다.

주제어 : 주식 매매 시스템, 주가 패턴, M&W 파동 패턴, 유전자 알고리즘, 전진 분석, 포트폴리오 최적화

## 제1장 서론

증권 거래가 시작된 이래 수많은 사람이 시장에서 초과이익을 얻기 위해 큰 노력을 기울여 왔다. 하지만 학자들 사이에 주가가 예측 가능하다는 주장과 예측 불가능하다는 주장이 첨예하게 맞서고 있다. Fama(1970)는 효율적 시장가설을 통해 주가는 다양한 정보가 즉각적으로 반영되어 움직이기 때문에 과거의 주가 자료를 이용하여 미래의 주가를 예측하는 것은 불가능하다고 하였다. 반면 Shiller(1981)는 효율적 시장가설을 비판하였는데 주식가격이 펀더멘탈(fundamental)로 간주하는 배당보다 심리적 요인 등에 의해 더 큰 영향을 받는다고 주장하였다. 군중행동(herding behavior)이 작용한다고 보았는데 투자자들이 낙관적인 분위기에 휩쓸려 자산시장에서 거품이 발생하기도 하고, 어떤 때는 비관적인 분위기에 동조하여 자산가격이 폭락하기도 한다고 하였다. 이런 Shiller의 주장은 2000년대 초 닷컴 버블(dot-com bubble) 붕괴를 예견한 것이 들어맞음으로써 명성을 얻었다. 또한, 주택시장 거품이 붕괴할 것을 경고하였는데 2008년 미국금융위기가 발생하면서 예측이 옳았음이 다시금 증명되었다.

효율적 시장가설을 주장하는 사람들은 기본적 분석뿐만 아니라 기술적 분석으로 초과 이익을 얻을 수 없다고 했으나 여전히 시장에서는 기본적 분석과 기술적 분석이 통용되고 있으며 남들보다 월등한 투자성적을 거두는 투자자들이 존재한다. 기본적 분석은 경제분석, 산업분석, 기업분석 등의 환경분석과 재무상태를 보는 재무분석으로 이루어져 있으며 기업의 내재가치보다 주가가 낮은 종목을 찾아 투자하려는 방법인 데 반해 기술적 분석은 모든 것을 차트에서 찾으려 한다. 기본적 분석보다 접근이 쉬워 주식에 투자하는 사람들 대부분은 기술적 분석에 관심이 더 많다.

기술적 분석의 방법도 여러 갈래로 나눌 수 있는데 가격지표, 추세지표, 거래량 지표, 시장지표 등의 기술적 지표 분석과 차트의 그래프 모양을 분석하여 추세를 예측하는 패턴 분석이 대표적이다. 기술적 지표분석 방법은 이동평균, MACD(Moving Average Convergence & Divergence), 스토캐스틱(stochastics), 볼린저 밴드(Bollinger band) 등 대중적으로 이미 많이 알려졌으며 연구 결과들도 많이 발표되었지만, 머리어깨형(head and shoulder pattern), 원형바닥형(rounding bottom pattern), 이중천장형(double top pattern), 이중바닥형(double bottom pattern), 깃대형(flags pattern), 삼각형(triangle pattern), 쐼지형(wedge pattern) 등의 패턴 분석에 관한 연구는 상대적으로 소홀했다. 패턴 분석으로 성공했다는 사람들은 책을 통해 자신의 기법을 주장하곤 하지만 사람의 눈으로 판단하는 가장 기초적인 접근법이다 보니 학술가치를 부여하기엔 부족하다.

최근에는 인공지능 분야에서 신경망을 비롯한 다양한 기계학습 기법을 사용한 빅데이터 분석이 금융시장에서 많이 활용되고 있다. 매일 쏟아지는 뉴스 기사와 공시 데이터를 이용하여 주가의 움직임을 예측하는 다양한 방법이 시도되고 있다. 또한, 이러한 기계학습 기법은 주가 차트의 패턴 연구에도 활용되고 있다. 주가 예측력이 있는 패턴들을 학습시켜서 매매 시점을 포착하려는 연구의 경우 가격의 단기 예측력은 높아졌지만, 장기 예측력은 한계가 있어서 장기 투자보다 단타 매매에서 활용되는 수준이다. 그리고 패턴을 얼마나 정확하게 찾아내느냐에 초점을 맞춘 연구도 있는데 찾아진 패턴이 매매에 적합한지 아닌지는 별개의 문제이기 때문에 실용적인 부분에서 취약할 수 있다.

본 연구에서 인용하고 있는 Merrill의 M&W 파동 패턴 분류 방법도 패

턴별로 발생 빈도와 성공 확률, 그리고 상승률과 하락률을 조사해서 가격 예측력이 높은 패턴을 찾는 게 목적이다. 모든 패턴을 전고점, 전저점이라 불리는 5개의 전환점으로 구분하는 단순함과 특정 패턴들에서 가격 예측력을 발견했다는 결과 보고가 있음에도 불구하고 실제 시장에서 활용한 성과 보고가 없었다. 이에 본 연구에서는 Merrill의 상승 전환, 하락 전환 32가지 패턴 분류를 시스템으로 구현하기 쉽게 10개 그룹으로 재해석하였으며 그룹마다 확률기반으로 1개 패턴만 선택해서 매매하는 방법을 제안한다. 특히 5점 패턴의 단순함은 패턴 인식의 정확도를 높이기 위해 드는 비용을 줄일 수 있는 장점이 있다. 확률기반의 선택 방법은 유전자 알고리즘을 이용하며 미래 수익에 대한 확신을 높이기 위해 다양한 기간의 전진 분석을 수행하였다.

본 논문은 다음과 같이 구성되어 있다. 2장에서는 본 연구의 주제인 M&W 파동 패턴과 최적화 방법으로 유전자 알고리즘 및 전진 분석에 관한 기존 연구를 검토하고자 한다. 3장에서는 패턴 전략의 성과를 측정하기 위한 시뮬레이터(simulator) 개발과 함께 실험 데이터의 선정에 관해 설명하고자 한다. 4장에서는 패턴 전략 시스템의 성과 분석을 종합적으로 정리하고 5장에서는 결론과 함께 앞으로 연구과제를 제시하겠다.

## 제2장 이론적 배경

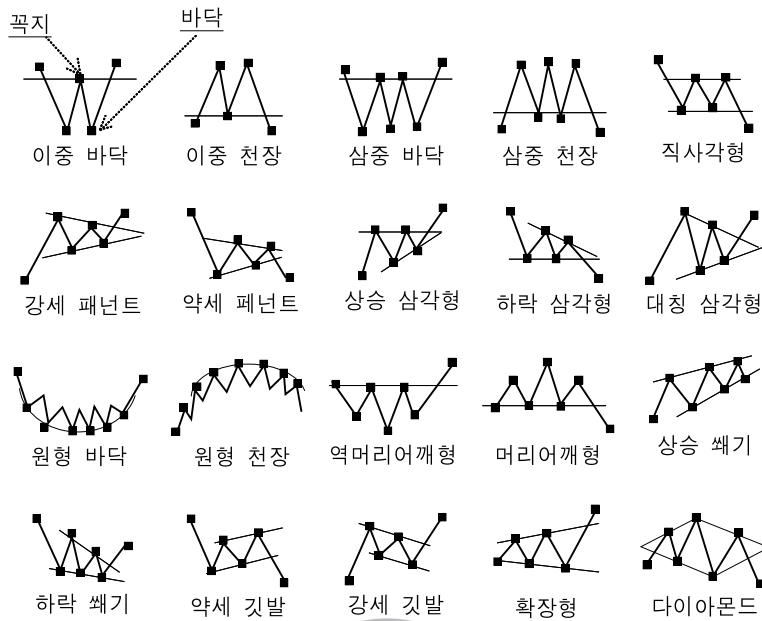
본 장에서는 주가예측을 위한 다양한 연구 방법들을 알아보고 패턴 분석과 연구 주제인 M&W 파동 패턴에 대해 고찰해 보겠다. 구현 과정에서 핵심인 유전자 알고리즘과 전진분석 기법에 대해서도 살펴보겠다.

### 2.1 주가예측을 위한 선행 연구들

많은 사람이 수익률을 높이기 위해 예측력이 좋은 주가예측 모델을 만드는 다양한 시도가 있었으며, 패턴 분석도 주가 예측을 위한 연구의 한 방향이다. 이 절에서는 패턴 분석 외에 어떤 연구들이 있었는지 국내 사례 중심으로 알아보겠다. 조지호(1989)는 증권투자수익률에 영향을 미치는 요인을 추출하여 증권투자수익률 예측모형을 만들었는데 경기변동별로 볼 때 불황기에 연평균 45%를 기록하여 회복기(15%)나 호황기(30%)보다 높게 나타났다고 하였다. 기업 규모별로는 불황기에 중기업이 53%를 보여 소기업(30%), 대기업(49%)보다 높았으며 회복기와 호황기엔 소기업이 중, 대기업을 앞질렀다고 하였다. 정용관과 윤영섭(1998)은 인공신경망(artificial neural network) 모형을 이용하여 주가예측 가능성을 검증하려고 하였으며 인공신경망 모형이 같은 구조의 선형모형보다 우월한 성과를 제공할 수 있다는 가능성을 발견했다고 하였다. 김해균과 김성신(2001)은 주가예측과 같은 비선형적인 시스템에서의 예측을 과거 데이터 학습을 이용한 신경망 기법이 아닌 퍼지 시스템(fuzzy system)의 도입으로 이루어 하였다. 주가지수의 이동평균선의 변형값을 입력 인자로 하는 퍼지 시스템을 실험하였으며 신경망보다 더 좋은 결과를 얻기도 하였다.

신동하, 최광호와 김복창(2017)은 신경망을 이용하여 기존의 보조지표에 대해서 다양한 조합으로 예측요소 적합 여부를 검토하고, 예측률 향상을 위해 최적의 보조지표 조합과 환율, 환율 이동평균, 전산업 생산지수 등 환경 예측요소들에 관해 연구하였다. 연구 결과 대부분 보조지표는 예측률을 저하하는 현상이 있었으며, 환경 예측요소들을 통해 예측률이 향상되었다고 하였다. 또한 RNN(recurrent neural network)과 LSTM(long short term memory network)이 DNN(deep neural network)에 비해 빠르고 안정적으로 학습했다고 하였다. 주가 등락이 심할수록 RNN과 LSTM이 DNN보다 예측 오차가 향상되었으며, 주가 등락이 적으면 예측 오차의 차이가 거의 없다고 하였다.

가격 예측에서 주식뿐만 아니라 상품자산의 가격을 예측하는 연구도 있었다. 송진호, 최홍식과 김전웅(2017)은 기계학습의 한 기법인 SVM(support vector machine)을 이용하여 골드만삭스 상품지수, 다우존스 USB 상품지수, 톰슨로이터 CRB 상품지수 등의 상품지수와 상품선물로 구성된 포트폴리오, 그리고 개별 상품선물을 대상으로 연구하였다. 주가지수, 무역지표, 고용지표, 경기 선행지표 등 19가지의 경제지표를 이용하여 상품지수와 상품선물의 등락을 예측하는 연구를 수행하였는데 결과적으로 상품지수에 투자하는 것보다 개별상품선물을 동일 비중으로 구성한 포트폴리오가 좋은 성과를 보였으며, 개별상품선물 각각을 예측하여 구성된 포트폴리오가 더 우수한 성과를 보였다고 하였다.



<그림 2> 차트 패턴

## 2.2 패턴 분석

윤재수(2017)는 증권 전문가들이 오랫동안 축적된 경험과 과거의 데이터 분석을 통하여 그래프의 모양에 따라 주가가 오르고 내린다는 사실을 발견하고 그래프의 모양을 여러 형태로 정형화하여 패턴이라는 것을 만들었으며, 이를 분석하여 앞으로 주가의 향배를 예측하게 되었는데, 이처럼 과거 주가 흐름을 통해서 검증된 패턴을 현재의 주가 흐름에 대입하고 주가의 등락을 예측하는 것을 패턴 분석이라고 하였다.

다수의 투자자가 매매 시점 판단의 근거로 사용하는 패턴 분석은 투자 분석 방법의 하나로 자리 잡고 있지만, 과거의 패턴이 미래에도 똑같이 되풀이되지 않는 경우가 많다는 단점이 있다. 그런데도 회자하는 유명한 패턴들로 이중 천장(쌍봉), 이중 바닥(쌍바닥), 원형 바닥, 머리어깨형 등

다수 존재하며 차트 분석을 다룬 책(한혜론, 2000; Bulkowski, 2006; 김정환, 2013; 윤재수, 2017)에는 빼놓지 않고 나오는 내용이다. 또한, 잘 알려진 캔들 차트 분석도 패턴 분석의 하나이다.

<그림 1>의 패턴의 종류를 자세히 들여다보면 법칙 내지는 원칙이라고 할 수 있는 규칙을 발견할 수 있다. 모든 패턴은 꼭지(peak, 전고점)와 바닥(valley, 전저점)이 지그재그(zig-zag)로 이루어져 있으며 이 꼭지와 바닥을 전환점(turning point)이라고 부른다. 이 전환점 중 고점끼리 또는 저점끼리 연결하여 방향성을 분석하는 것이 추세선 분석이다(김정환, 2013). 일부 트레이딩 툴에서 차트상에 추세선을 보여 주고 있지만, 그래픽 처리이고 실제 매매에 활용할 수 없다. 사용자들의 요구에 비해 전산화가 덜 된 분야로 여겨진다.

Kamijo and Tanigawa(1990)는 **주가 패턴** 인식에 RNN을 적용하여 삼각형 패턴을 인식하도록 하였으며 패턴을 정확히 찾는 것에 목적을 두었다. 백진우와 조성준(2000)은 **종합주가 지수** 데이터에서 머리어깨모형 중 왼쪽 어깨에 해당하는 데이터로 신경망을 학습시킨 후 이를 이용하여 현재 혹은 테스트 데이터를 입력으로 주어 성능을 평가하였다. Leigh et al.(2002)는 지식공학(knowledge engineering)과 데이터 마이닝(data mining)을 이용하여 강세 깃발(bull flags) 패턴을 찾으려 하였다. 10 X 10 격자에 120일(다른 연구자료에서는 60일)간의 데이터를 10%씩 잘라서 배치하는 방법으로 강세 깃발형을 인식하도록 하였으며 NYSE Composite Index에 적용하여 효과가 있음을 입증하였다. 국내에서는 이재영과 한치근(2003)이 같은 방법으로 60일 데이터를 한 단위로 하여 5가지의 변형된 매수 패턴과 역시 5개의 매도 패턴을 정의했으며 KOSPI200



지수에 적용하여 매수 패턴 발생 후 20일 후 청산, 매도 패턴 발생 후 20일 후 청산하는 방식으로 기대 수익률을 계산하였다.

이종우 외(2003)는 특정 시계열 패턴에 부합하는 주식을 추출하였고 추출된 주식 집합에 여과 규칙을 적용하여 실제 매매 대상이 되는 주식을 골라내었다. 시계열 패턴은 기술 지표 값들의 조합으로 표현되며, 독특한 분포를 보이는 구간에 존재하는 기술 지표 값들이 주가의 향후 움직임에 영향을 준다는 가정하에 여과 규칙은 독특한 분포를 보이는 구간 내의 데이터 값들로부터 자동으로 유도되었다. 시뮬레이션을 통해 이러한 방법이 시장 평균 수익률을 웃돈다는 사실을 확인하였다.

Liu and Kwong(2007)은 웨이블릿 분석(wavelet analysis)과 Multi resolution analysis(MRA) 등을 활용하여 주식 시계열 데이터 기반에 가능한 불규칙성의 인식 과정을 자동화하는 알고리즘을 제시하였는데 <그림 1>의 차트 패턴들을 얼마나 정확하게 찾는지에 목적을 두었다. 이재원과 오장민(2007)은 거래 시스템의 최종 성능을 개선하기 위해 신경망의 예측 결과와 다양한 미시 거래 정책 인자들을 통합적으로 고려하여 다중의 시뮬레이션(simulation)을 수행하는 ‘통합 다중 시뮬레이션(integrated multiple simulation) 기법’을 제안하였으며 기술적 분석에서 사용되는 여러 가지 파동 패턴 중, V자형 패턴이 매우 유용한 패턴임을 확인하였다.

Bao and Yang(2008)은 시계열 데이터에서 전환점을 인식하고 기술적 지표 30여 가지를 조합해서 확률 모델로 매매 전략을 학습시켜서 무작위성과 불확실성을 감소시키도록 하였다.

최홍식, 김선웅과 박성철(2011)은 시스템트레이딩 전략과 패턴인식 알고

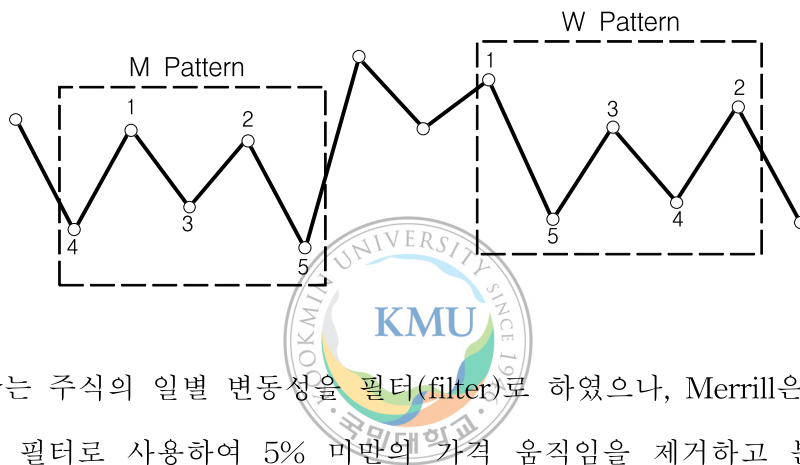
리즘인 CHMM(continuous hidden markov model)을 융합한 KOSPI200 트레이딩시스템을 만들어 일반적인 시스템트레이딩의 성과, 시스템트레이딩 전략 없이 패턴인식 알고리즘만을 사용했을 때의 성과와 비교하였다. 그 결과 시스템트레이딩 전략과 패턴인식 알고리즘을 융합한 시스템이 가장 많은 수익을 안정적으로 내고 있음을 확인하였다.

이재원(2012)은 일반적인 차트 패턴 모형 외에 주가의 단기적 변동을 반영하는 이동 평균선 패턴과 장기적인 주가 변동 과정을 표현하는 전환점 행렬(turning point matrix) 방식을 제안하였다. 이재원(2013)의 연구에서는 고변동성 주가 패턴들을 독립적으로 학습하여 거래에 활용하는 패턴 기반의 주식 거래 시스템을 구축하였다. 유재필 외(2017)는 클러스터링(clustering)을 통해서 패턴을 분류하였고, 데이터 시각화를 위해 방사형 차트를 활용하여 그 결과를 분석하였다.

이모세와 안현철(2018)은 그래프를 입력 값으로 사용하여 주식시장의 방향(상승 또는 하락)을 예측하는 이진 분류기로써 합성곱신경망(convolutional neural network, CNN)을 적용하였다. 주식시장에서 사용하는 주요 기술지표들의 5일간의 변화 양상을 그래프로 만들어 그 그래프를 입력 데이터로 사용하였다. 실제 데이터에 로지스틱 회귀분석, 인공신경망, SVM과 같은 비교모형을 적용하여 예측 정확도를 비교하였으며 통계적으로 유의한 차이가 있다고 하였다.

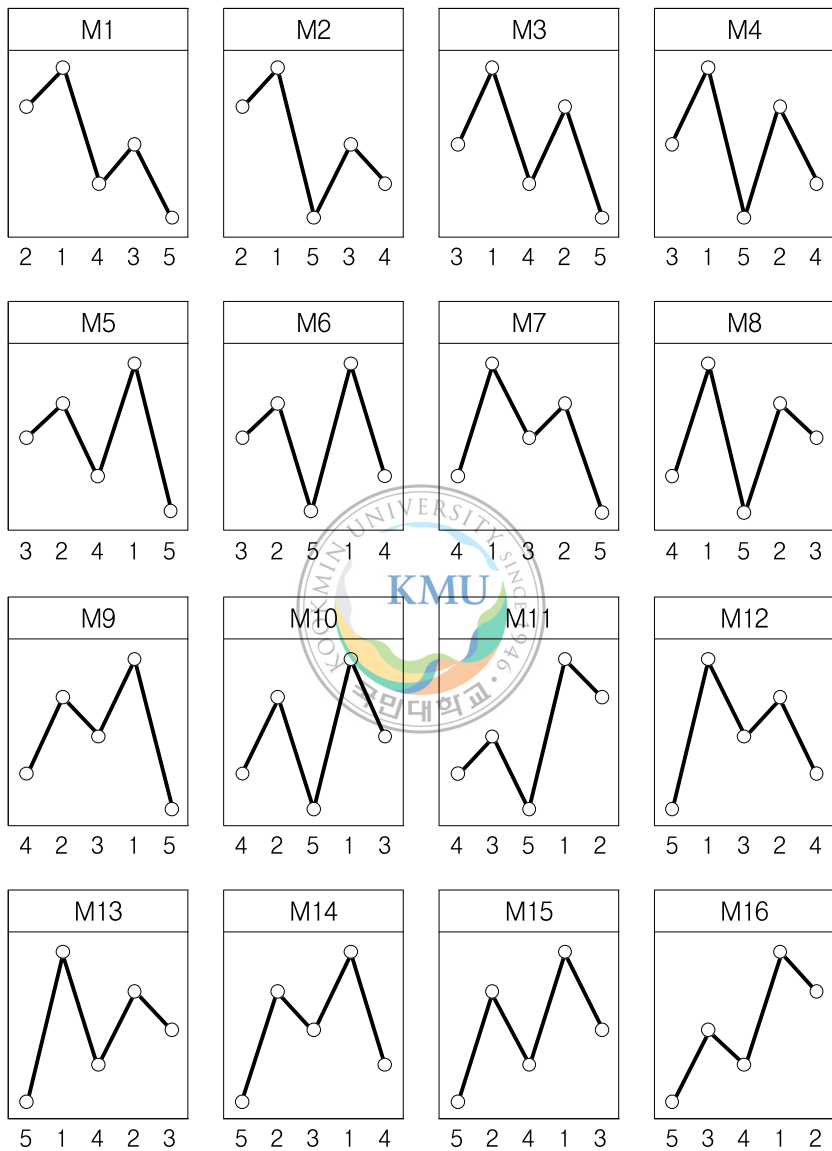
## 2.3 M&W 파동 패턴

Levy(1971)는 5개 전환점만으로 일련의 패턴을 정의하였다. 32가지의 패턴을 찾아냈으나 아쉽게도 예측 가치를 찾는 것은 실패하였다. Merrill(1980)이 Levy의 생각을 이어받아 5점 패턴을 완성하였으며 상당한 예측력이 있음을 발견하였다.

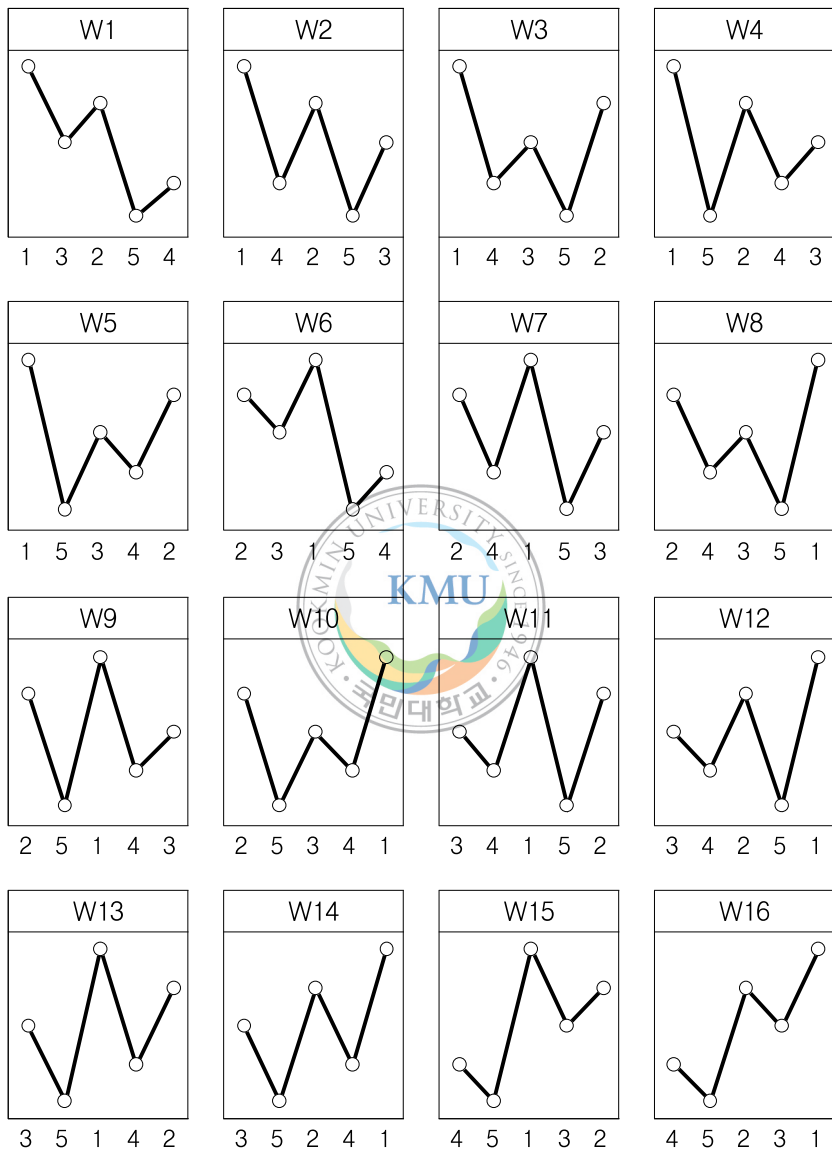


Levy는 주식의 일별 변동성을 필터(filter)로 하였으나, Merrill은 5% 변동률을 필터로 사용하여 5% 미만의 가격 움직임을 제거하고 분석했으며,<sup>1)</sup> Levy의 32가지 패턴을 <그림 2>에서와 같이 M자 16개와 <그림 3>과같이 W자 16개로 분류하였다.

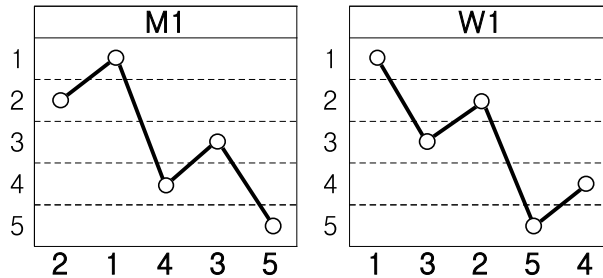
1) 직전의 고점에서 5% 이상 가격이 하락하면 추세가 바뀌었다고 보고 신규 저점을 인식하는 방법으로 가격이 5% 이상 하락하지 않으면 마지막 전환점은 고점인 상태를 유지한다. 반대로 저점 이후 5% 이상 가격이 상승해야 다음 고점을 인식한다.



<그림 2> Merrill의 M자 파동 패턴



<그림 3> Merrill의 W자 파동 패턴

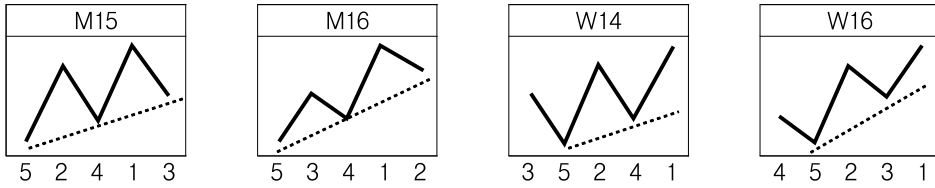


각 패턴의 하단에 붙어 있는 번호는 전환점의 높이 순서를 나타낸다. 예를 들어 M1형에서 2 1 4 3 5는 두 번째 전환점의 높이가 가장 높고 다섯 번째 전환점의 높이가 가장 낮다. W1은 첫 번째 전환점이 가장 높고 세 번째 전환점이 2순위이고 두 번째 전환점이 3순위다. Merrill은 M자 패턴이 나오면 상승으로 보았고 W자 모양이 나오면 하락으로 보았다. 하지만 강세장과 약세장에서는 반대로 진행되기도 한다고 하였다. 강세장에서 M자 패턴이 나오면 하락의 시작으로 볼 수 있다고 하였다.

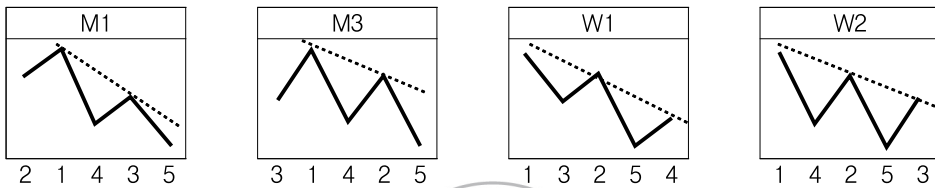
1898년부터 다우존스 산업에서 5% 필터를 적용한 전환점을 모두 찾아 분석했으며 패턴마다 발생 빈도와 패턴 발생 후 평균 상승률, 평균 하락률을 측정하여 패턴의 유용성을 보여주었다. 1984년에 추가로 발표한 자료에서는 상승 확률과 하락 확률을 다루었다. M자 패턴 16개 중 12개가 상승 확률이 50% 이상이었고 W자 패턴도 16개 중 13개 패턴의 상승 확률이 50%를 넘었다.

Nyemera(n.d)은 전통적인 차트 패턴을 M자와 W자 패턴으로 설명하였다.

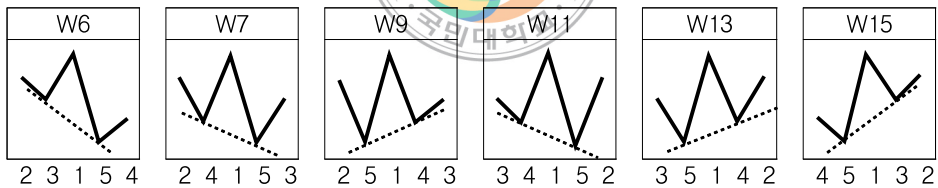
상승추세 : M15, M16, W14, W16



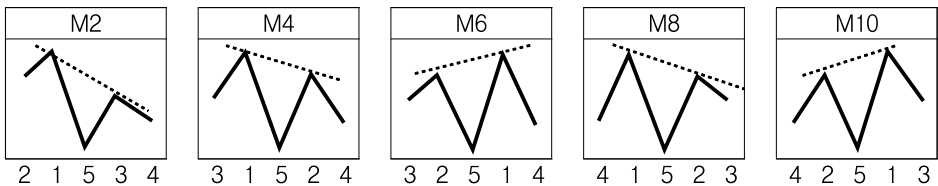
하락추세 : M1, M3, W1 및 W2



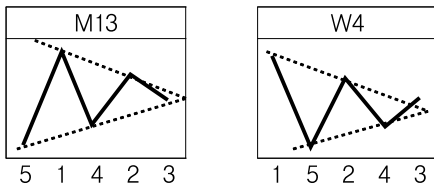
머리어깨형 : W6, W7, W9, W11, W13 및 W15



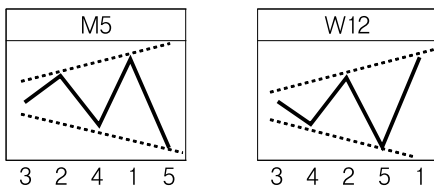
역머리어깨형 : M2, M4, M6, M8 및 M10



삼각형 : M13 및 W4

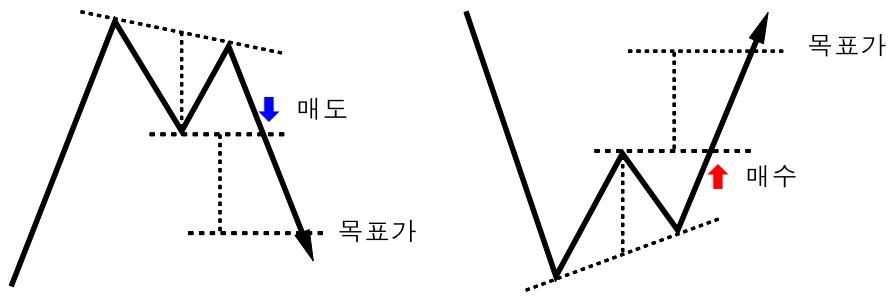


확장 패턴 : M5 및 W12



이렇듯 M&W 파동의 패턴으로 웬만한 패턴의 설명이 가능하다. 그리고 M&W 패턴을 인터넷에서 검색하면 이중 천장, 이중 바닥의 개념으로 설명한 자료도 나오는데(SirForex.net) Merrill의 그것과는 완전히 다른 것이다<그림 4>. 여기에서 M자 패턴은 이전 저점을 이탈하는 경우 매도하고 W자 패턴은 이전 고점을 돌파할 때 매수하는 방식이다. 매도 및 매수 이후 목표가 계산 방식도 제시하고 있는데 매매 방식이 비교적 단순해서 학술적인 연구자료는 못 찾았다.





<그림 4> 이중 천장, 이중 바닥의 M&W 패턴

## 2.4 유전자 알고리즘

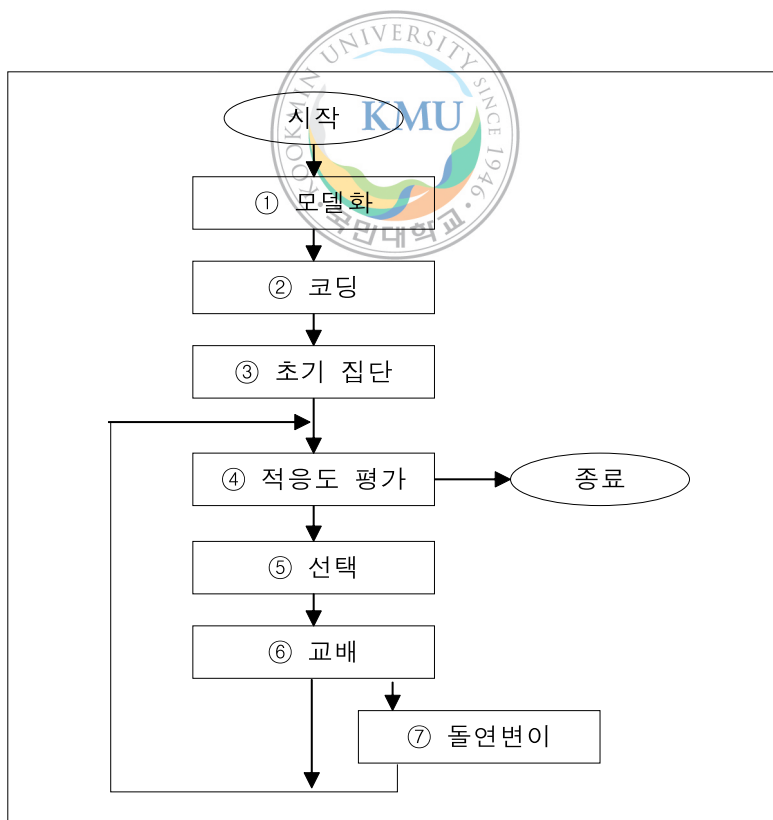
GA는 자연계의 진화과정을 모사한 확률기반 계산 알고리즘으로서 최적화 문제와 관련한 연구들에서 가장 많이 사용되고 있는 인공지능 알고리즘 중 하나다. GA는 John Holland에 의해 창안되었으며, 교차(crossover), 돌연변이(mutation), 그리고 역치(inversion)와 같이 유전학에 따라 고안된 연산자들과 함께 일종의 “자연의 선택”을 사용하여 “염색체”(예를 들어 비트(bit), 즉 1과 0의 문자열)의 한 개체집단에서 새로운 집단으로 이동시키는 방법이다(Mitchell, 1996).

GA의 처리 순서는 대상 문제의 특징을 추출하여 목표 상태를 정의하고 (모델화) 유전자(기호열)와 적응도 평가 함수를 정의한(코딩(coding)) 다음 초기에 적당한 유전자를 갖는 개체를 생성하여(초기 집단) 적응도를 평가하고 목표 상태가 되면 종료한다(적응도 평가). 적응도가 가장 낮은 유전자의 개체를 가장 적응도가 높은 유전자로 대체하여 적응도 낮은 유전자의 개체는 버리고 적응도가 높은 개체는 중복되게 선택한다(선택). 개체 간에 유전자를 교환하고(교배) 다시 적응도 평가단계로 넘어간다. 필요

에 따라서는 적응도와 관계없이 새로운 유전자를 끼워 넣는데(돌연변이) 랜덤(random)하게 유전자를 구성하여 국소해(local minimal)로 빠지는 것을 막는다(아사이 노보루, 2017).

GA는 최적화, 자동 프로그래밍, 기계학습, 경제학 등 다양한 분야에서 응용되고 있으며 매개변수를 많이 사용하는 시스템 매개에서도 전수 테스트를 못 하고 근사치로 최적화할 때 사용되는 기법이다.

조희연과 김영민(2003)은 유전자 알고리즘을 통해 기존의 대표적인 기술적 투자기법인 MACD 지표와 스토캐스틱 지표를 구성하고 이를 통해 초과수익이 가능한지 파악하였다. 유전자 알고리즘에 의한 수익률과 구매-유지 전략과 무위험 투자인 국고채 수익률과 비교하였다. 실험 결과는



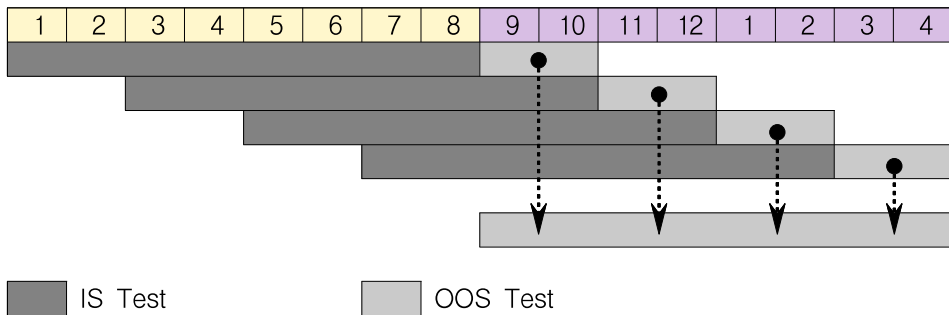
<그림 5> 유전자 알고리즘 처리 순서도

비교 대안 투자 수익률보다 우수하였다고 하였으나 거래 비용을 포함하지 않아 과대 평가되는 문제점이 있다. 이주상 외(2004)는 유전자 알고리즘을 이동평균기법에서 가장 높은 수익률을 얻는 해 찾기에 적용하였다. 이동평균선을 이용한 매매기법은 대세 상승기에는 장기 보유 전략보다 우월한 해를 구할 수 없었고 등락을 거듭할 때 상대적으로 좋은 해를 구할 수 있었다고 하였다.

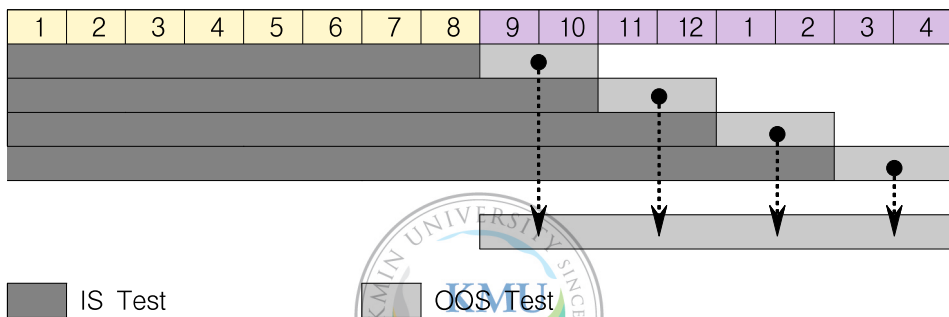
김명종(2009)은 SVM 앙상블 학습의 최적화 문제를 해결하기 위해 GA를 사용했으며, 김선웅과 안현철(2010)은 주가지수 등락의 예측 모형으로 SVM을 사용하는 지능형 시스템 연구에서 이중 임계치(two-threshold)를 최적화하기 위해 GA를 사용하였다. 이종식과 안현철(2017)은 주가지수 추세를 상승추세, 박스권, 하락추세의 다분류(multiple classification) 체계로 확장하여 주가지수를 예측하고자 다분류 Support Vector Machines(multiclass SVM, MSVM)을 사용하면서 성능 향상을 위한 래퍼(wrapper)로서 GA를 이용하였다.

김은총과 오경주(2019)는 은닉 마르코프 모델(hidden Markov model; HMM)을 통해 개별 종목들의 국면을 파악하고 투자 대상을 선정하였으며, GA를 활용하여 종목의 투자 비중 최적화를 통해 높은 투자성과를 얻을 수 있는지 실증분석하였다. HMM으로 종목을 선택하여 동일 비중 포트폴리오를 구성했을 때와 개별 종목의 투자비중을 최적화했을 때를 비교하였는데 샤프 비율(Sharp ratio)이 유의미하게 증가하였음을 확인하였다.

강성석, 이현준과 오경주(2019)는 펀드투자 의사결정 지원모형에서 GA를 활용하였다. 제시한 일부 모델에 대해서는 GA로 투자 비중을 최적화한 것이 동일 비중 투자방식보다 성과가 안 좋았다고 하였는데 이는 전체



(a) 회전방식 WFA



(b) 고정방식 WFA

<그림 6> 전진분석 방법

리 과정과 알고리즘 간 융합이 수반되어야 함을 의미한다고 하였다.

## 2.5 전진 분석

다수의 시스템 매매를 하는 사람들은 시스템의 수익성을 검증하기 위해 과거 데이터로 최적화하는 과거 테스트(back test)의 결과에 의존하는 경향이 있다. 과거 테스트에서 수익이 없다면 미래의 수익은 더욱 기대할 수 없으므로 과거 테스트의 결과는 중요하다. 따라서 약간의 수정과 최적화 작업을 통해 과거 테스트 결과를 높이는 데 열중하곤 한다. 하지만 미

래의 주가가 과거와 같게 움직인다는 보장이 없으므로 결국은 과최적화로 귀결되는 경우가 많다. WFA는 Pardo(1992)가 창안했으며, 과거 테스트만으로 의사결정 할 때 저지르게 되는 과최적화의 실수를 줄여주는 강력한 방법이다. 방법은 매매 전략을 표본 내 데이터(in-sample data, IS)로 최적화한 다음 그 매개변수를 그대로 표본 외 데이터(out-of-sample data, OOS)에 적용해 봄으로써 미래에도 여전히 신뢰할 만한 성과를 낼 수 있는지 검증한다. WFA는 한 구간의 테스트가 끝나면 바로 다음 구간을 테스트하는 회전방식(rolling type)과 테스트의 시작을 고정해 놓은 채 전체 구간을 늘려가면서 반복하는 고정방식(anchored type) 등이 있다<그림 6>.

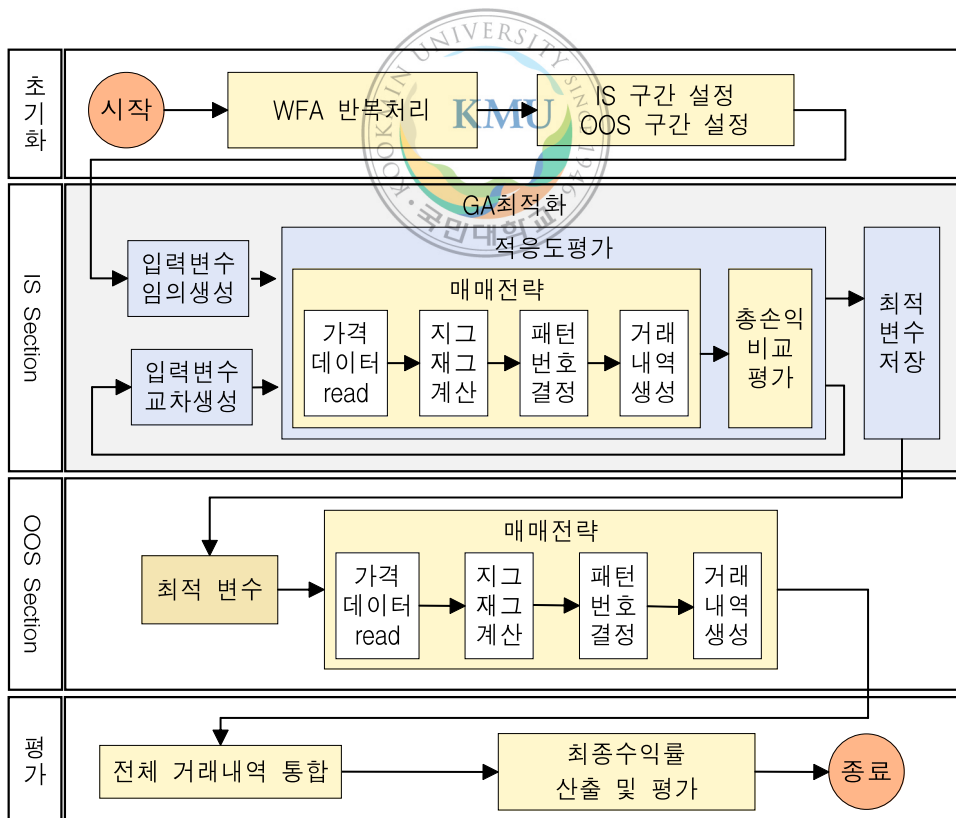
Pardo는 WFA의 첫 번째 장점이 전략의 실제 매매 능력을 검증하는 것이라고 하였고, 두 번째 장점은 최적화 후 이익과 위험비율에 대해 정확하고 신뢰할만한 척도라고 했으며 매매 전략이 엄격한 WFA를 통과하면 실제 매매가 가능하다고 하였다.

WFA는 매매 전략의 최적화뿐만 아니라 AI 기법에서도 최적화의 한 방법으로 사용하고 있다. Cao and Tay(2003)는 적응형 매개변수가 있는 SVM과 표준 SVM 비교 실험에서 WFA 루틴(routine)을 이용하여 전체 데이터 세트를 5개의 학습(training) - 확인(validation) - 테스트(testing)의 세트로 나누어 실험하였다. Zbikowski(2015)도 SVM 분류기로 단기 추세를 예측하는 검증 실험에서 WFA를 사용하였다.

### 제 3 장 시스템 개발

본 연구에서는 종목별로 시가, 고가, 저가, 종가 가격 데이터를 날짜순으로 읽어 가면서 M자 패턴이 나오면 사고, W자 패턴이 나오면 파는 것을 기본 전략으로 한다. 적용할 패턴과 버릴 패턴의 조합을 GA 최적화로 찾았으며 WFA를 통해 시장이 변해도 최적화를 통해 지속해서 수익을 내는지 점검하였다.

파이썬(Python)으로 개발하였으며, 시뮬레이터의 처리순서는 WFA 모듈(module)에서 GA 모듈을 실행시키고 GA 모듈 중 적응도 평가 단계에서 매매전략에 해당하는 모듈을 실행시키게 되어 있다<그림 7>.



<그림 7> 시뮬레이터 구성도

### 3.1 매매전략 수립

2장 이론적 배경에서 나왔던 패턴 연구 중 대부분은 정교한 패턴 인식에 할애하고 있다. Kamijo와 Tanigawa(1990)은 삼각형 패턴 하나만 그리고, Leigh et al.(2002)는 강세 깃발 하나만 인식하려고 하였으며 Liu와 Kwong(2007)은 확장형, 썰기형, 삼각형, 이중 바닥, 이중 천장, 머리어깨형, 역머리어깨형까지 인식하는 진일보한 방식을 제안했다. Leigh et al.(2002)는 강세 깃발 후 얼마나 오르고 내리는지를 조사했으나 Kamijo와 Tanigawa(1990) 그리고 Liu와 Kwong(2007)은 패턴 인식에 초점을 두고 있다. 이는 숲은 못 보고 나무만 보고 있는 형국이다. 패턴을 정확하게 찾아내는 것이 진정한 의미가 있으려면 해당 패턴이 이론대로 움직여 주어야 한다. 예를 들어 삼각 패턴을 찾았다 하더라도 예측한 방향으로 움직일 거란 보장이 없으므로 찾은 패턴이 매매에 적합한지 아닌지 확인하는 작업이 추가되어야 한다.

Merrill(1980)의 M&W 파동 패턴 연구는 주가 예측에 무게를 두고 있지만, 패턴 분류와 패턴별 예상 수익률만으로는 매매 전략을 만들 수 없다. M1 패턴 이후 기하평균 상승률이 13.89%였고 M13은 12.79%, M6는 11.86%, M4는 11.12%, M15는 11.15%를 평균적으로 상승했다고 하였다. W자 패턴 이후 하락률은 W2, W4, W7, W9이 11% 이상 하락률을 보였다고 하였다. 이렇듯 몇 가지 패턴에서는 신뢰할 만한 수준의 결과를 발견했기 때문에 확률적으로 유리한 패턴을 조합하면 될 것 같으나 측정하는 시점에 따라 값이 다르게 나오기 때문에 주기적으로 재조사를 해야만 실제 매매에 적용할 수 있을 것이다. Merrill도 패턴별 상승 및 하락의 순위는 흥미로울 수 있지만, 결정적인 것으로 간주해서는 안 된다고 하였다.

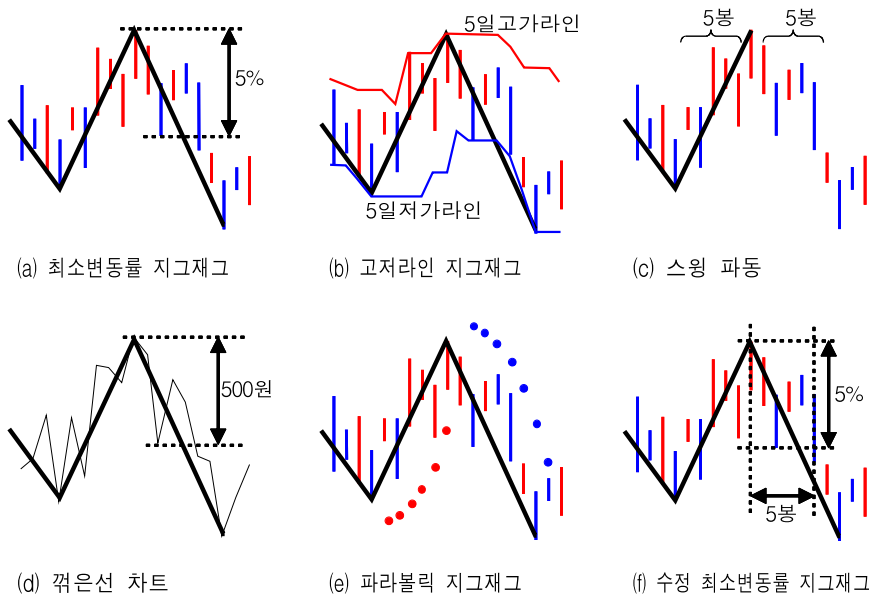
1984년에 추가로 발표한 자료에서는 상승 확률과 하락 확률을 다루었는데 M자 패턴 16개 중 6개의 상승 확률이 60% 이상이었고, 1개가 하락 확률이 59% 이상이었다. 반면 W자 패턴은 16개 중 4개 패턴의 상승 확률이 60%를 넘었고, 하락 확률이 60% 이상인 패턴은 없었다. W자 패턴 이후 하락이 뚜렷하지 않아 보인다.

과거 패턴이 미래에도 같게 움직이지 않는다고 하지만 시장의 추세는 순간적으로 바뀌지 않고 얼마간 유지되는 것을 경험상 느끼고 있다. 측정할 수 없지만, 김원조(2016) 외 다수의 주식 투자자들이 주가에도 관성의 법칙이 적용된다고 말한다. 그래서 기술적 분석을 통한 매매가 가능하고 또한 패턴을 이용한 매매도 가능하다. 예를 들면 M1 패턴이 발생하고 나서 대부분 14% 상승했다면 언제까지일지는 모르겠지만, 당분간은 M1 패턴 이후에 14% 상승을 기대할 만하다는 것이다.

본 연구에서는 Merrill의 M&W 파동 패턴이 KOSPI 시장을 얼마나 잘 설명하는지를 증명하고자 하는 것이 아니다. 2장에서 Nyemera(n.d)는 주가 차트에 나오는 대부분의 패턴 모형을 M&W 파동 패턴으로 설명하였다. 거창하게 여길 수도 있으나 32개의 패턴은 알파벳에 비유할 수 있고 이들 패턴을 이용하여 매매전략을 구사하는 것은 알파벳을 이용하여 단어를 표현하는 것에 비유할 수 있다. 아무렇게나 나열된 알파벳은 의미도 없고 단어가 될 수 없다. 의미가 있는 즉 수익이 발생하도록 패턴들을 조합하는 것이 본 연구의 방향이다.

주가 데이터를 가지고 M&W 파동 패턴을 인식하려면 먼저 전환점의 가격과 위치를 계산한 다음 전환점들의 높낮이와 순서를 비교해서 해당 패턴을 찾는다. 앞장에서 Kamijo와 Tanigawa(1990)은 RNN 기법을,





<그림 8> 전환점 계산 방식의 종류

Leigh et al.(2002)는 격자에 데이터를 대입해서, Liu와 Kwong(2007)은 웨이블릿 분석과 MRA 등의 기법을 적용하여 패턴을 인식하려 했음을 설명했다. 이들 방식의 난이도에 비하면 Merrill(1980)의 방식은 비교적 간단하다고 말할 수 있겠다.

Merrill은 1980년 자료에서 5% 변동률 필터를 사용하였고 1984년 자료에서 10% 필터를 사용하였다. <그림 8>에서 (a) 최소변동률 지그재그(minimum change rate zig-zag) 파동이 이 방법이며 엘리엇 파동 이론(Elliott Wave Principle)에서 파동을 셀 때도 이 방법이 사용되었다(Hill, 2010). 전환점을 판별하는 방법은 이외에도 몇 가지 더 있는데 인터넷에서 아이디어 차원에서 논의되는 수준이라 계산 방식에 상관없이 그냥 지그재그 파동 또는 지그재그 지표로 통용되고 있다. 따라서 다른 사람들은 <그림 8>의 지그재그 파동을 다른 이름으로 부를 수도 있다. 다만 c) 스

윙파동(Swing wave)은 많이 알려진 방식이고 지그재그 파동과 구별된다.

계산방식을 바꾸면 전환점의 가격과 위치가 조금씩 다르게 산출되기 때문에 매매 결과에도 영향을 미친다. 본 연구에서는  $n$ 일 고점과  $n$ 일 저점을 지그재그로 잇는 (b) 고저라인 지그재그(high-low line zig-zag) 방식과 좌우  $n$ 개의 봉보다 가운데 봉의 가격이 가장 높거나 낮을 때를 고점과 저점으로 인식하는 (c) 스윙 파동 방식을 추가로 테스트하였다.

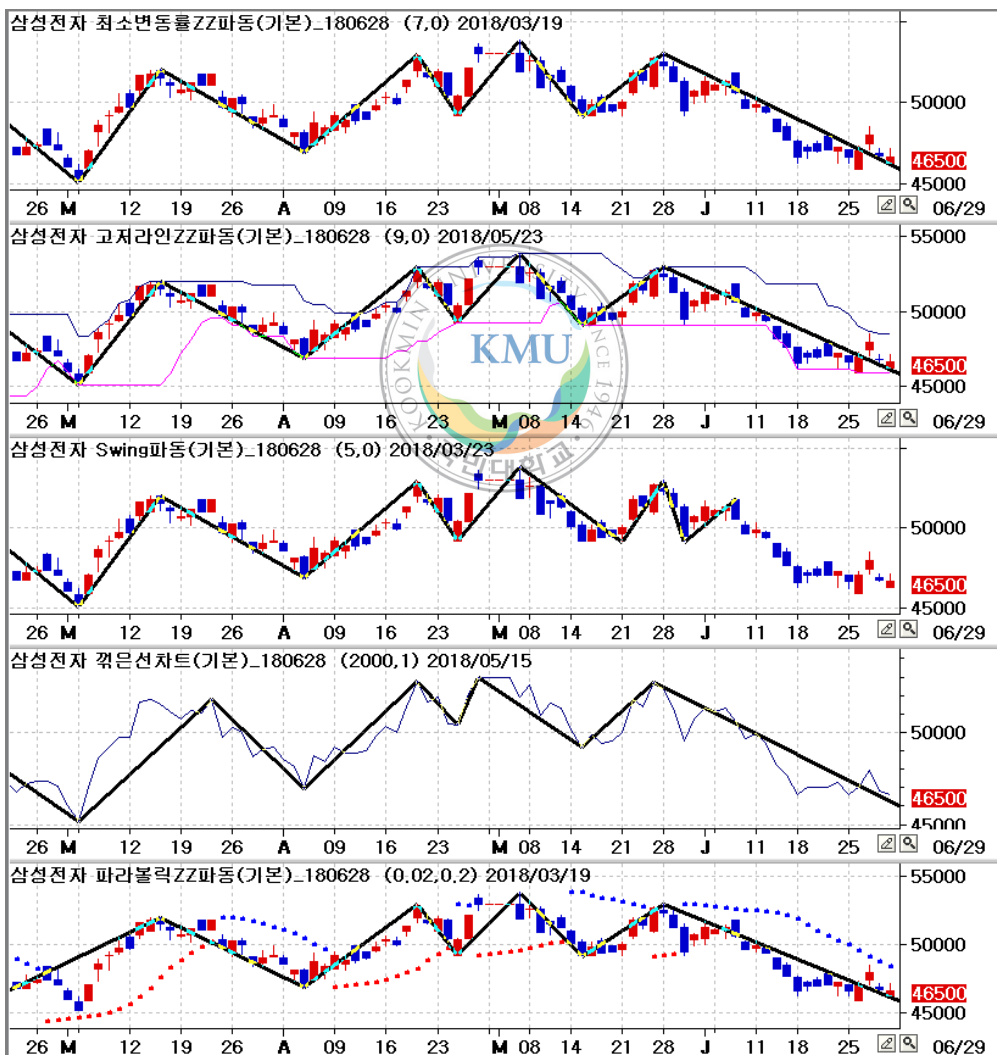
(d) 꺾은선 차트 방식은 최소변동률 지그재그 방식과 거의 모든 것이 같다. 하나 다른 것이 있다면 변동률이 아닌 변동 폭 즉 가격 차이를 기준으로 고점, 저점을 인식한다. 따라서 단일 종목을 분석하는 경우는 문제 없지만, 가격대가 다른 여러 종목을 동시에 분석하기엔 적합하지 않다. 예를 들어 현재가가 5,000원인 주식에서 500원의 변동 폭과 50,000원인 주식에서의 500원의 움직임은 완전히 다르기 때문이다. 본 연구에서는 20종목으로 구성된 포트폴리오의 성과를 측정하였기 때문에 꺾은선 차트 방식은 배제하였다. 그러나 비슷한 가격대의 종목끼리 포트폴리오를 구성하는 경우에는 꺾은선 차트 방식도 괜찮은 방법일 수 있다.

(e) 파라볼릭 지그재그 방식은 Wilder(1978)가 만든 파라볼릭 SAR (Parabolic Stop And Reverse) 지표를 이용하여 고점, 저점을 인식한다. 주가가 SAR보다 아래 있는 동안은 하락 추세이므로 가장 낮은 가격을 저점으로 인식하고 주가가 SAR 보다 위에 있는 동안은 상승 추세이므로 가장 높은 가격을 고점으로 인식한다. 그 외는 다른 지그재그 파동 계산과 같다.

(f) 수정 최소변동률 지그재그는 Bao와 Yang(2008)이 전환점을 계산할 때 사용한 방식으로 따로 이름을 붙이지 않았지만 다른 방식과 구분하기

위해 편의상 이름을 붙여 놓았다. 최소변동률 지그재그에서 변동률과 함께 변동 기간도 조건으로 추가한 방식이다. 최소변동률을 5%, 최소 기간을 5일로 하면 5% 이상 변동이 있어도 5일이 지나지 않았다면 전환점 판단을 않는다.

5가지 지그재그 계산방식은 원리가 서로 달라 같은 구간에 대해 고점과

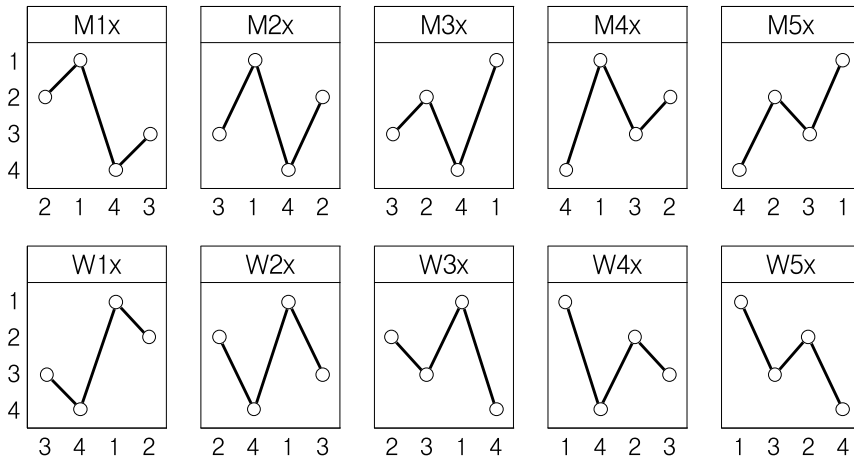


<그림 9> 지그재그 파동의 실제 적용 예

저점이 다르게 계산되는데 그 특성을 알아 둘 필요가 있다. 최소변동을 지그재그는 가격 즉 수직적으로 측정하기 때문에 가격 움직임이 적은 횡보장에서 휩소(whipsaw, 거짓 속임수 신호)를 걸러주지만 시장 불안으로 단기간에 가격이 널뛰기할 때는 휩소를 고스란히 보여준다. 기간까지 고려된 수정 최소변동을 지그재그는 단기 휩소 구간을 피해가게 해준다. 고저라인 지그재그는 기간 개념 즉 수평적으로 측정되기 때문에 가격 변동이 적은 구간에서도 일정하게 전환점을 찾아주는 편이다. 스윙 파동은 우측의 봉 개수가 조건을 만족해야 전환점을 인식하기 때문에 우측의 봉 개수만큼 늦게 반응을 하게 된다.

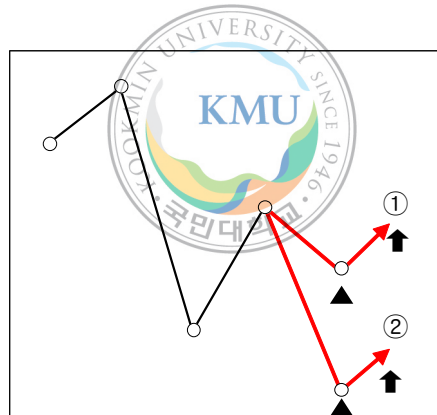
최소변동을 지그재그 또는 고저라인 지그재그를 사용하여 전환점을 계산하는 경우 예를 들어 M자 패턴을 인식했을 때 매수 조건 충족으로 보고 바로 매수 진입을 할 것이다. 가격이 반등에 성공했다면 저점에서 싸게 매수하였기 때문에 수익률이 높을 것이지만 저점이 더 내려갈 수도 있는 위험이 내포되어 있다. 스윙 파동으로 계산하는 경우에는 하락하던 저가가 더는 내려가지 않고 n일 동안 신규 저가가 나오지 않아야 M자 패턴으로 인식한다. 다른 지그재그 방식보다 높은 가격에서 매수 진입하기 때문에 수익률 면에서 불리하지만 추세 전환을 확인하고 진입하기 때문에 위험을 조금이라도 줄일 수 있다.

32개 패턴이 모두 유효한 것이 아니므로 매매에 이용하려면 수익이 나는 패턴을 선별해야 한다. 본 연구에서는 Merrill의 패턴 분류를 재해석하여 비슷한 패턴끼리 묶은 후 매매 알고리즘을 설계하였다. 먼저, 4개의 전환점으로 표현할 수 있는 패턴을 찾아보면 총 10개이다<그림 10>. 편의상 첫 번째 패턴을 M1x라고 명명하였으며 Merrill의 분류 번호와 무관하

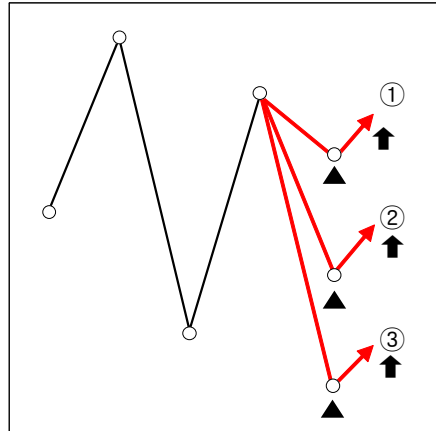


<그림 10> 4점 패턴

다.

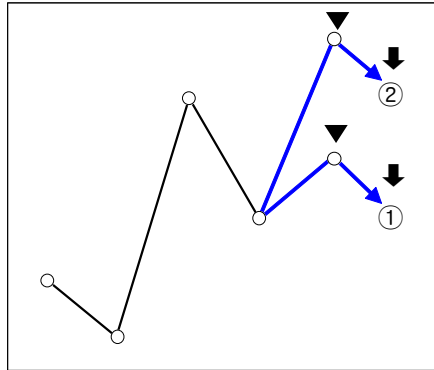


M1x 이후 전개 상황은 위와 같이 ①에서 반등할 수도 있고 ②에서 반등할 수도 있다. 어느 시점에 진입하는 것이 더 유리한지 알 수 없으므로 과거 데이터를 통해 성공 확률이 높은 시점을 찾는 것이 핵심이다. 진입 시점의 구분을 위해 ①의 상태를 M11로 정의하였고, ②의 상태를 M12로 정의하였다.

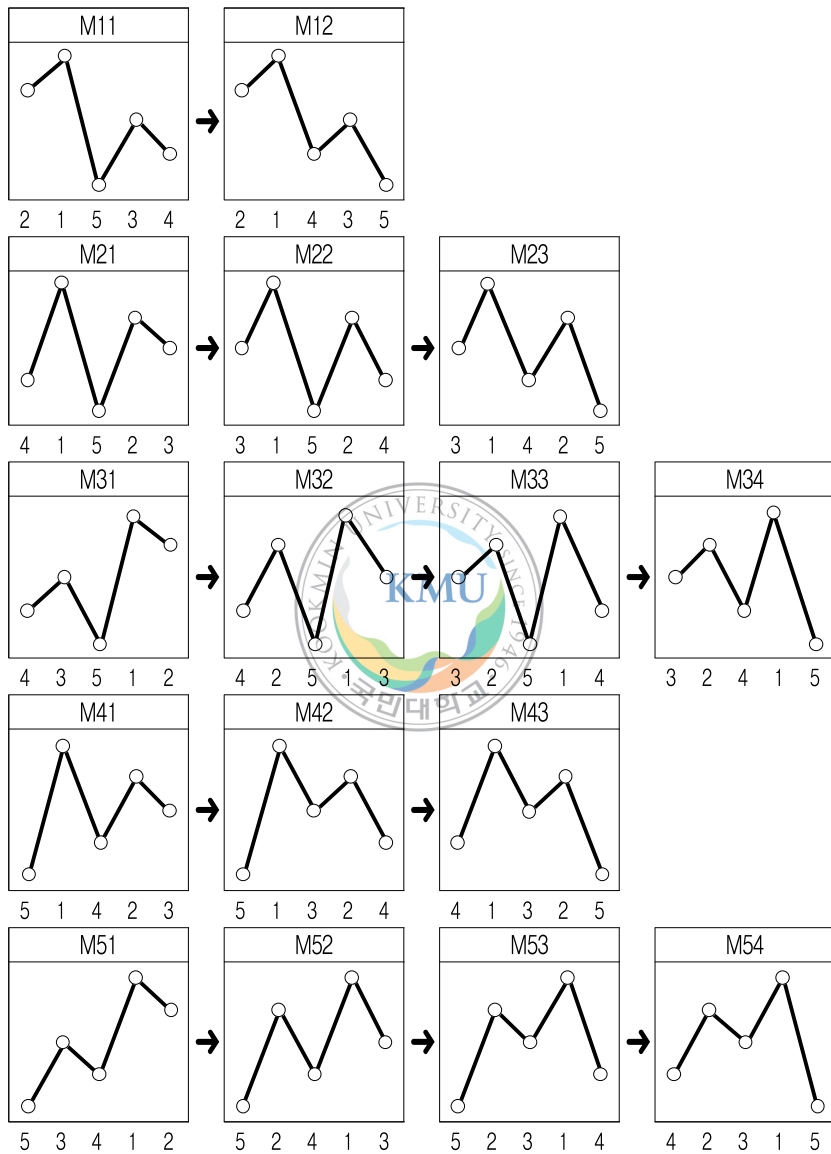


다른 예로 현재 가격 차트의 상태가 <그림 10>의 M2x이고 가격이 하락 중이라면 예상되는 시나리오는 3가지가 있다. 어느 시점에 진입할지는 직전 기간의 시뮬레이션 결과에 따라 정해진다. 위 그림에서 ①의 상태를 M21, ②의 상태를 M22, ③의 상태를 M23으로 정의하였으며, <그림 10>의 10가지 경우에 대해 같은 방식으로 번호를 부여한 것이 <그림 11>과 <그림 12>이다.

패턴 정의가 완료되면 패턴이 진행 중인 상태에서 진입하는 것이 유리한지 패턴이 완성된 다음 진입하는 것이 더 나은지 검토가 필요하다. 패턴이 완성되기 전에 매수 진입하면 저점에서 매수하기 때문에 수익 면에서 유리할 수 있지만, 추가 하락의 위험이 크다. 반대로 패턴의 완성을 확인하고 진입하면 승률이 높은 대신 수익이 줄어든다. 위 그림에서 ‘▲’ 표시는 패턴이 진행 중인 상태에서 진입하는 것이고 ‘↑’ 표시는 패턴이 완성된 다음 진입하는 경우이다. 최소변동률 지그재그 방식 또는 고저라인 지그재그 방식을 적용하면 ‘▲’ 지점에서 패턴을 인지하고 진입할 수 있다. 스윙 파동 방식을 사용하면 ‘↑’ 지점 이후에 패턴을 판별할 수 있다.

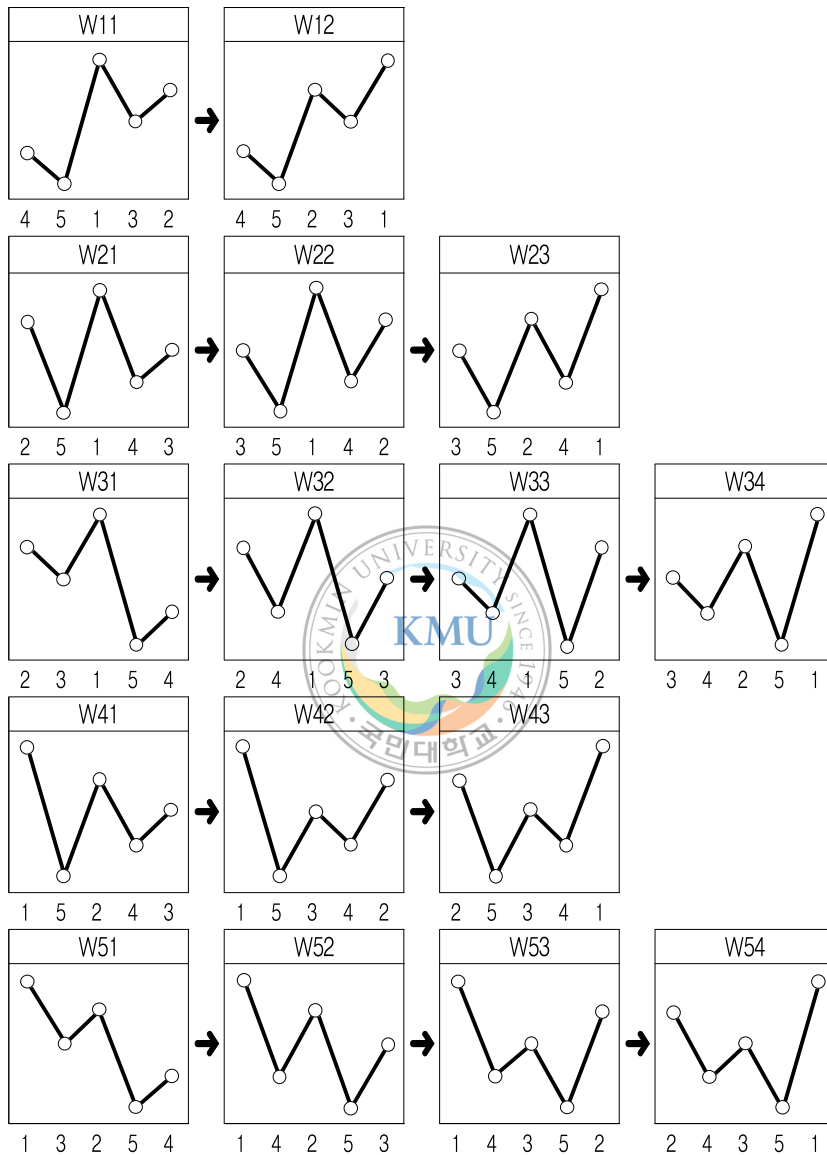


청산 시점을 찾는 방법도 진입 시점을 찾는 방법과 같다. 위 그림은 W1x 상황에서 가격 상승을 보여주고 있다. ①은 전고점을 못 뚫고 반락하는 경우이고 ②는 전고점을 뚫은 다음 반락하는 경우이다. M자 패턴과 마찬가지로 ①의 상태를 W11로 정의하였고 ②의 상태를 W12로 정의하였다. W1x 패턴이 과거 기록에서 전고점을 뚫고 올라가는 경우가 많았다면 시스템은 W11에서 청산하지 않고 기다렸다가 W12가 되면 청산할 것이다. 만일, W12가 되기 전에 가격이 하락한다면 청산 기회를 놓친 셈이고 다음번 W자 패턴이 나올 때까지 포지션(position)은 계속 유지된다.

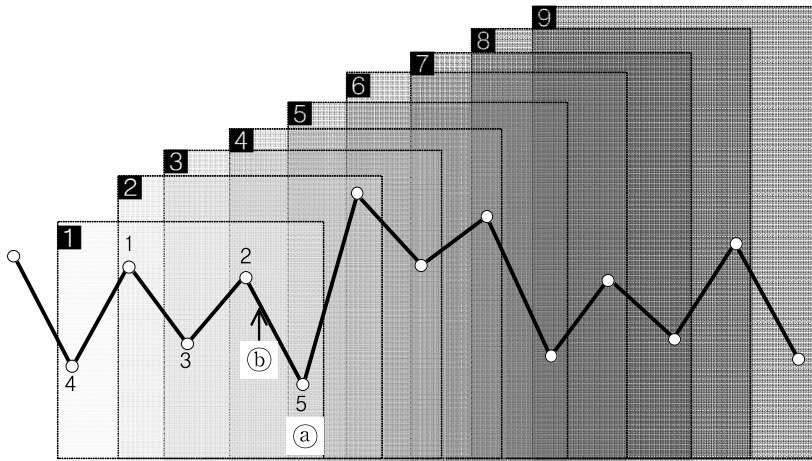


<그림 11> M자 패턴의 재분류





<그림 12> W자 패턴의 재분류



2장 이론적 배경에서 사용된 파동 그림에 새로운 분류 번호를 대입하면 다음과 같다.

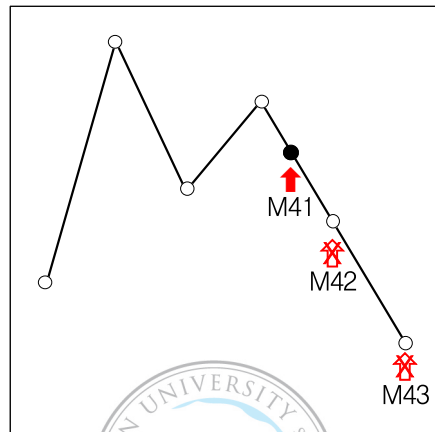

**1**M43 → **2**W54 → **3**M31 → **4**W11 → **5**M42 → **6**W51 → **7**M11 → **8**W42 → **9**M54

처음에 나오는 M43 패턴이 완성되는 시점인 ①에서 매수하고자 한다면 매매 로직을 다음과 같이 생각해 볼 수 있다.

if (M4패턴번호 == 1 or M4패턴번호 == 2 or M4패턴번호 == 3) then  
Buy()

하지만 의도와 다르게 위 로직으로는 ② 지점에서 매수 신호가 발생한다. 아래 그림처럼 가격은 M41 모양 다음에 M42가 오고 마지막으로

M43의 상태가 된다. M41이 가장 먼저 오기 때문에 매수 신호는 항상 M41에서만 나오게 된다. M41 패턴에서 매수가 발생하면 M42와 M43 상태에서는 이미 매수 포지션이기 때문에 매수 신호가 발생할 수 없다. 그러므로 각 그룹에서 취할 수 있는 패턴 번호는 하나만 존재해야 한다.



M43 상태에서 매수를 발생시키려면 다음과 같이 M4x 그룹에서는 한가지 번호만 있는 코드를 만들어야 한다.

```
if (M4패턴번호 == 3) then
```

```
    Buy()
```

M4x 그룹이 아닌 M31과 M11, M54 등과 같이 다른 그룹의 패턴은 M43에 영향을 미치지 않기 때문에 다음과 같이 코드를 작성할 수 있다.


```
if (M1패턴번호 == 1 or M2패턴번호 == 3 or M3패턴번호 == 1 or
```

```
    M4패턴번호 == 3 or M5패턴번호 == 4) then
```

```
    Buy()
```

이상의 개념 정의를 이용하여 작성한 것이 <그림 13>의 슈도코드(pseudocode)이다. 첫 번째 줄에서 in\_1부터 in\_10까지 10개의 매개변수(parameter)는 <그림 10>의 10가지 경우의 수를 의미한다. 현재의 가격 상태에서 지그재그 방식으로 5개 전환점의 가격을 계산하고 순서를 매기면 패턴 번호를 부여할 수 있다. 전고점 이후 가격이 하락하고 있는 상태에서 전환점의 높이가 3 1 5 2 4로 나오면 변수 m2x의 값은 2가 되며 M22를 의미한다.

나머지 변수 m1x와 m3x, m4x, m5x의 값은 0이다. 패턴 번호가 결정되고 나면 입력 매개변수와 비교한다. m2x의 값과 in\_2의 값이 일치하면 매수가 발생하고 그렇지 않으면 다음 가격을 읽어서 반복 처리한다. 마지막



```

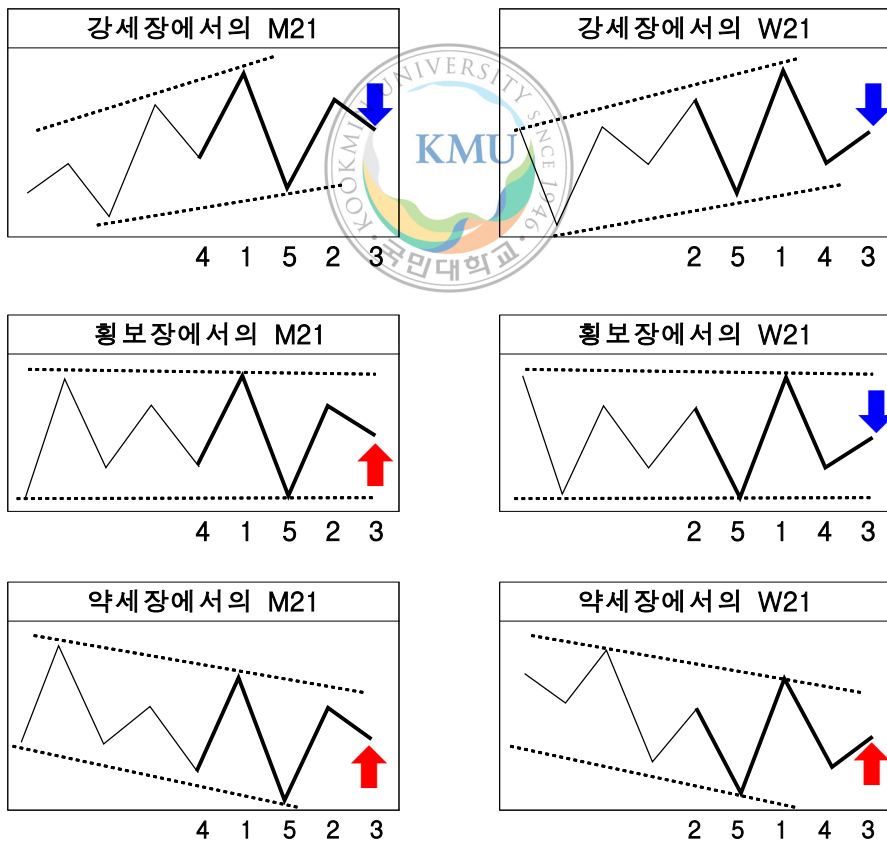
def trading(in_1, in_2, in_3, in_4, in_5, in_6, in_7, in_8, in_9, in_10)
    m1x = 0
    if (전환점높이 == 21534) then m1x = 1
    else if (전환점높이 == 21435) then m1x = 2
    m2x = 0
    if (전환점높이 == 41523) then m2x = 1
    else if (전환점높이 == 31524) then m2x = 2
    else if (전환점높이 == 31425) then m2x = 3
    ... ..
    if (m1x == in_1 or m2x == in_2 or m3x == in_3 or m4x == in_4 or
        m5x == in_5) then
        Buy()
    if (w1x == in_6 or w2x == in_7 or w3x == in_8 or w4x == in_9 or
        w5x == in_10) then
        Sell()

trading(1, 1, 3, 2, 3, 2, 1, 3, 3, 4)
    
```

<그림 13> 매매전략 기본 로직

줄의 trading() 괄호 안에 10개의 숫자가 나열되어 있다. 앞의 5개는 M자 패턴의 번호를 뒤의 5개는 W자 패턴의 번호를 의미한다. <그림 13>의 예제는 M11, M21, M33, M42, M53의 패턴이 나오면 매수하고 W12, W21, W33, W43, W54 패턴이 나오면 매도하는 매매 전략이다. 괄호 안의 매개 변수의 값이 바뀔에 따라 매매 결과는 다르게 나올 것이고 이 숫자들을 최적화하는 것이 본 연구의 핵심이다.

2장에서 기본적으로 M&W 과동 패턴은 반전형이지만 강세장이나 약세장에서는 반대일 수 있다고 하였다. 강세장에서의 M자 패턴과 약세장에서



<그림 14> 시장추세를 반영한 패턴 매매

서의 M자 패턴, 그리고 횡보장에서의 M자 패턴의 결과가 다르다. Merrill의 조사 결과에는 강세장과 약세장이 언급되어 있으나 어떤 상황을 강세장으로 보고 어떤 상황을 약세장으로 구분하였는지 설명이 나와 있지 않다.

본 연구에서는 주가가 계속 오르고 있으면 강세장이고 하락하고 있으면 약세장, 나머지는 횡보장으로 정의하였다. <그림 14>에서 보는 바와 같이 M자와 W자 패턴의 1번 고점이 직전의 전고점 중 가장 높고 5번 저점보다 낮은 전저점이 있으면 강세장으로 정의했고 5번 저점이 가장 낮은 저점이고 1번 고점보다 높은 고점이 존재하면 약세장으로 정의하였다. 그 외는 횡보장으로 보았다. M자 패턴이면서 약세장 또는 횡보장이면 매수 처리하고 강세장이면 매도 처리하였다. W자 패턴은 강세장과 횡보장에서 매도 처리하고 약세장일 때 반대로 매수하였다. 이렇게 패턴 분류 번호뿐



<그림 15> M&W 파동 패턴을 이용한 매매

만 장상황까지 일치해야 매수와 매도가 가능하도록 하였으며 매개변수는 10개에서 30개로 늘었다. <그림 15>는 매매 알고리즘의 검증을 위해 상용 도구인 예스트레이더(YesTrader)에 적용해 본 결과이다.

<표 1> 사용 변수 목록

변수명	기본	최소	최대	증가	변수 설명
변동률	4	2	10	1	지그재그 계산에서 최소 변동의 크기
손절률	40	0	60	5	진입가에서 입력 값 이상으로 손실 발생하면 청산
목표청산	0	0	60	5	진입가에서 입력 값 이상으로 이익 달성되면 청산
최소수익률	0	0	60	5	최소목표치
수익감소율	0	0	60	3	최소목표 이후 입력 값 이상으로 하락하면 청산
본전청산	0	0	60	5	진입가에서 입력 값 이상으로 이익 발생했다가 본전이 되면 청산
고점대비	0	0	60	5	진입 후 고가 대비 입력 값 이상으로 하락하면 청산
in_1	1	0	2	1	강세장에서 M11~M12이면 청산
in_2	1	0	3	1	강세장에서 M21~M23이면 청산
in_3	1	0	4	1	강세장에서 M31~M34이면 청산
in_4	1	0	3	1	강세장에서 M41~M43이면 청산
in_5	1	0	4	1	강세장에서 M51~M54이면 청산
in_6	2	0	2	1	약세장에서 M11~M12이면 진입
in_7	3	0	3	1	약세장에서 M21~M23이면 진입
...	...	...	...	...	...
in_27	4	0	4	1	횡보장에서 W31~W34이면 청산
in_29	3	0	3	1	횡보장에서 W41~W43이면 청산
in_30	4	0	4	1	횡보장에서 W51~W54이면 청산

## 3.2 유전자 알고리즘 최적화

### 1단계. 초기집단

본 연구에서 유전자 개체는 8개로 하였으며 M&W 파동 패턴 전략에 사용되는 변수와 강제 청산 전략에 사용되는 변수까지 최적화 대상으로 하였다. 변수마다 값의 범위가 넓지 않기 때문에 8바이트(byte)의 2진수로 변환하여 사용하였다.

증가단위가 1보다 큰 변수는 증가단위로 나눈 다음 정숫값을 이용했다. 변수에 들어갈 수 있는 값을 0에서 60까지 임의의 정수라고 정하면 특정 숫자가 할당될 확률은  $1.6\%(=1/61)$ 가 된다. 반면 증가단위를 5로 지정하면 변수에 들어갈 수 있는 숫자는 0, 5, 10, 15, ..., 55, 60이 되어 특정 숫자가 할당될 확률은  $7.7\%(=1/13)$ 가 된다. 정교한 값을 얻기 위해서는 1단위가 맞겠으나 오히려 최적화 후 결과는 5단위로 하는 것이 더 좋았다. 유전자 알고리즘을 사용하는 것 자체가 어렵값을 찾는 것이기 때문에 값을 세분화하면 오히려 부정적인 효과가 커진다.

```
[[0 0 0 0 1 0 0 1 0 0 0 0 0 1 1 1 0 0 0 0 1 0 0 0 ... 0 0 0 0 0 0 1 0 0]
[0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 ... 0 0 0 0 0 0 0 0]
[0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 1 1 0 0 0 0 1 0 1 0 ... 0 0 0 0 0 0 1 1]
[0 0 0 0 0 0 0 1 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 1 ... 0 0 0 0 0 0 1 0]
[0 0 0 0 0 1 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 1 0 1 ... 0 0 0 0 0 0 1 1]
[0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 ... 0 0 0 0 0 0 1 0]
[0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 1 1 0 0 0 0 0 0 1 1 1 ... 0 0 0 0 0 0 1 1]
[0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 1 1 0 0 0 0 0 1 0 0 ... 0 0 0 0 0 0 0 0]]
```

<그림 16> 초기집단 설정 결과 예시



## 2단계. 적응도 평가

8개의 유전자 개체를 순서대로 10진수로 환원하여 입력변수 목록을 만든 다음 앞서 설명한 매매전략으로 시물레이션을 수행한다. 20종목의 시물레이션이 완료되면 총손익을 합산하여 번호대로 저장하고 순위를 매긴다.

총손익만으로 순위를 판단하는 경우 1~2종목의 수익률이 매우 높고 나머지 종목이 모두 손실이 나더라도 1순위가 될 수 있다. 이러한 경우가 과최적화이며 다음 OOS 구간에 거래가 발생하지 않을 수 있다. 그러므로 일부 종목의 매매 결과에 편중되지 않도록 포트폴리오 구성 종목 수의 2/3에서 매매가 발생하는지 확인하는 로직을 반영하였다.

## 3단계. 선택

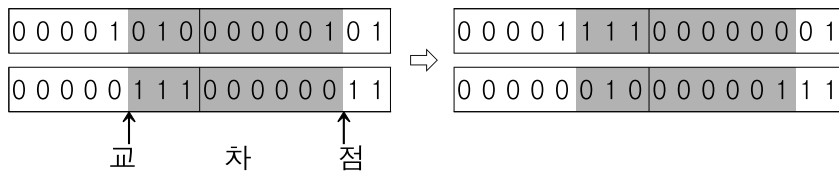
세대교체에서 유전자 선택법은 엘리트(elite) 보존 방식을 적용하였다. 1순위 유전자의 형질이 더 많이 살아남도록 1순위 유전자를 꼴찌 유전자에 복제하여 1순위 유전자 개체를 두 개로 만든다. 높은 순위 유전자와 낮은 순위 유전자 2개씩 조합하여 부모 쌍을 4개 만들고, 부모 쌍으로 유전자 교환을 수행한다.

수익률1 = trading(2, 35, 35, 50, 24, 50, 25, 1, 2, 1, 0, ..., 2, 4, 0, 1)  
수익률2 = trading(9, 20, 40, 20, 0, 20, 20, 1, 3, 2, 2, ..., 2, 0, 3, 0)  
수익률3 = trading(4, 40, 0, 0, 0, 0, 0, 1, 1, 1, 1, ..., 3, 4, 3, 4)  
....  
수익률8 = trading(8, 20, 20, 35, 12, 25, 60, 2, 0, 3, 0, ..., 0, 1, 0, 3)

<그림 17> 적응도 평가 과정 예시

#### 4단계. 교차

유전자 교차 방식은 여러 방식이 있는데 본 연구에서는 여러 개의 유전자를 일제히 교환하는 다점 교차 방식을 사용하였으며 교차점은 3개로 하였다. 교차 후 각 변수는 새로운 값을 가지는데 2진수로 구성되어 있으므로 일부는 변수의 범위를 벗어나는 경우가 있다. 이때 해당 구간만 난수를 발생시켜 범위 안으로 들어오도록 조정한다.



<그림 18> 다점 교차 방식

<그림 18>의 2진수 값을 10진수로 변환하면 <그림 19>와 같다.

변수 A	변수 B	변수 A	변수 B
10	5	15	1
7	3	2	7

<그림 19> 10진수 변환 예시

변수 A의 범위를 0~10으로, 변수 B의 범위를 1~5로 지정했다면 교차 후에 1번의 변수 A는 15, 2번의 변수 B는 7로서 지정 범위를 벗어나므로 한 번 더 난수 처리한다.

#### 5단계. 돌연변이

적응도 평가, 선택, 교차를 반복 수행하다 보면 어느 순간 8개 개체 모두 유전자의 값이 같아지는 경향이 있는데 국소해로 빠질 우려가 있으므로

로 될 수 있으면 돌연변이를 많이 수행하였다. 여섯 번에 한 번꼴로 돌연변이를 생성했는데 뒤로 갈수록 돌연변이의 개수를 많이 넣었다.

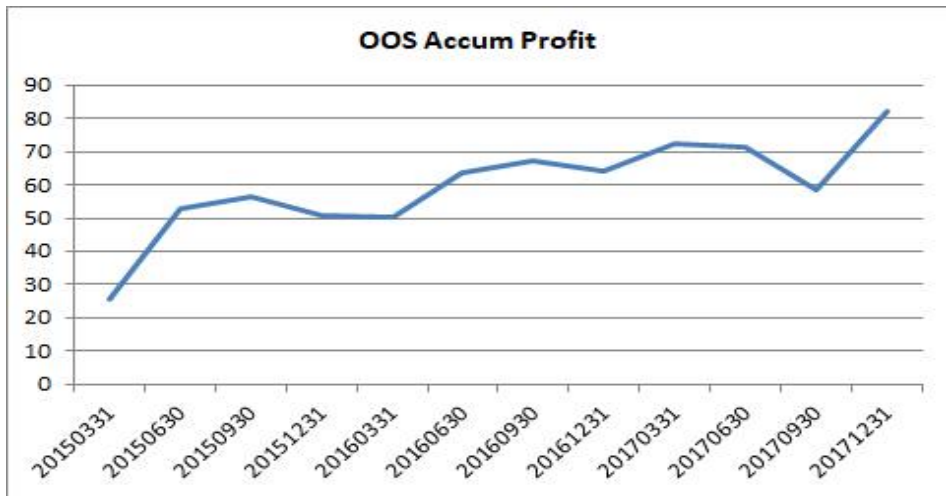
### 3.3 전진 분석 적용

WFA에서 IS와 OOS의 기간을 어떻게 설정하느냐에 따라 전체 수익의 결과도 많은 차이를 보인다. 현실적으로 테스트를 무한정 할 수 없으므로 전체 수행 시간을 고려하여 <표 2>처럼 12가지로 테스트해 보았다.

IS 구간에 GA로 찾은 최적 변수를 OOS 구간에 적용하여 나온 매매 결과들을 연결하여 최종 수익률과 최대손실률 등을 계산하였다. OOS 구간의 예상 수익을 정확히 계산하기 위해 매 OOS 구간 종료일 다음 날의 시가에 청산하도록 처리하였다.

<표 2> WFA 테스트 케이스

번호	방식	IS 기간	OOS 기간
1	회전식	2 년	3 개월
2	회전식	2 년	6 개월
3	회전식	3 년	3 개월
4	회전식	3 년	6 개월
5	회전식	1 년	3 개월
6	회전식	1 년	6 개월
7	고정식	3 년	3 개월
8	고정식	3 년	6 개월
9	회전식	2 년	1 개월
10	회전식	2 년	12 개월
11	고정식	3 년	1 개월
12	고정식	3 년	12 개월



<그림 20> OOS 누적수익률

M&W 패턴에 의해서만 진입, 청산이 되도록 시뮬레이터를 만들면 OOS 구간에서 매매가 없는 경우도 발생한다. IS 구간에서 매수 신호가 발생해서 그대로 매수 포지션을 유지했으면 수익이 발생했을 상황인데 구

<표 3> 거래내역

날짜	매매종류	신호명	수량	가격	거래금액	수익(원)	수익률
2016-07-01	매수	B	1,904	5,250	-9,926,980	0	0
2016-07-28	매수청산	XL0	1,904	5,570	10,533,214	606,234	5.0
2016-12-20	매수	B	2,724	3,670	-10,030,694	0	0
2017-01-02	매수청산	XL1	2,724	3,620	9,797,956	-232,739	-2.3
2017-01-02	매수	B	2,762	3,620	-9,935,981	0	0
2017-03-31	매수청산	XL0	2,762	3,410	9,357,518	-578,463	-6.7
2017-04-17	매수	B	2,958	3,380	-10,032,826	0	0
2017-07-03	매수청산	XL1	2,958	3,630	10,669,062	636,236	6.4
2017-07-03	매수	B	2,754	3,630	-9,933,265	0	0
2017-07-07	매수청산	XL0	2,754	3,625	9,919,564	-13,701	-1.1

간을 잘라서 테스트하는 바람에 기회를 놓치게 된다. 그래서 IS 구간 종료일까지 매수 포지션을 유지하고 있으면 그다음 이어지는 OOS 구간 시작일의 시가에 재진입하도록 처리하였다. <표 3>의 거래내역을 보면 매수 청산과 매수 진입이 동시에 발생한 구간이 두 군데 있다. 이는 M&W 패턴의 매매 신호와 무관하게 OOS 구간 종료 시점에 청산하고 IS 구간의 매수 포지션을 이어받아 재진입한 것이다.

### 3.4 테스트 대상 종목군 선정

본 연구는 주식 포트폴리오 시스템에 적용하고자 최소 20종목에 대해 최적화를 실시하였다. 개별 종목마다 따로 변수 최적화를 하면 과최적화의 우려가 크므로 과최적화를 피하면서 분산투자의 효과를 높이기 위해 구성 종목 수를 20으로 선정하였다. 종목 수를 늘리면 GA와 WFA 테스트에 걸리는 시간도 따라서 증가하기 때문에 성능적인 측면도 고려하였다.

종목 선정 기준은 시가총액, 일간 변동성, KOSPI 지수와의 베타(Beta) 3가지로 하였다. 각 기준에 따라 상위 20종목, 하위 20종목을 추출하여 총 6개 포트폴리오로 나누어 테스트하였으며 포트폴리오 간에 종목이 겹치는 것은 인정하였다. 우선주는 제외하고 보통주만을 대상으로 순위를 매겼으며 표 5에서 시가총액은 한국거래소 홈페이지의 2014년 12월 29일 자 자료를 참조하였다. 종목명은 현재 시점의 종목명으로 바꿨다.

시가총액을 기준으로 종목을 선정할 때 현시점을 기준으로 시가총액 상위 종목을 선택하면 실험을 하기도 전에 오류에 빠진다. 이미 성과가 좋

은 종목들이기 때문에 웬만한 전략으로 과거 테스트를 하더라도 어느 정도 수익이 발생하게 되어 있다. 현재의 시총 상위 종목들이 미래에도 시총 상위에 머물러 있을 거란 보장을 못 한다. 그래서 매매 시작을 2015년 1월 2일부터로 정하고 그 전날 기준으로 종목을 선택하였다. 변동성은 2012년 1월부터 2014년 12월까지 3년 동안의 일간 변동성을 계산한 다음  $\sqrt{252}$ 를 곱하여 연환산하였다. 베타 역시 2012년부터 2014년까지 3년 동안의 베타 값이다. 베타 값은 아래 공식에 의거 산출하였다.

$$B_S^M = \frac{\sum_{i=1}^n (R_S^i - \overline{R_S})(R_M^i - \overline{R_M})}{\sum_{i=1}^n (R_M^i - \overline{R_M})^2}$$

$$= \frac{n \sum_{i=1}^n (R_S^i \cdot R_M^i) - \sum_{i=1}^n R_S^i \cdot \sum_{i=1}^n R_M^i}{n \sum_{i=1}^n (R_M^i)^2 - (\sum_{i=1}^n R_M^i)^2}$$

Where  $R_S^i$  := 주식종목의 일간로그수익률

$R_M^i$  := 시장의 일간로그수익률, 여기서는 KOSPI지수

$\overline{R_S}$  := 주식종목 로그수익률의 평균

$\overline{R_M}$  := 시장 로그수익률의 평균

### 3.5 시뮬레이션 환경 설정

매매 수수료율은 0.2%로 하였고 청산 시 세금은 0.3%로 계산하였다. 슬리피지(Slippage)<sup>2)</sup>는 진입과 청산 각각에 대해 1 호가로 하였다. 초기 투입자금은 2억 원으로 설정하였으며 1종목당 진입금액은 1천만 원으로 고정하였다.

### 3.6 거래 내역 통합 및 재가공

IS 구간에서 GA를 이용해 최적화하는 과정은 1종목씩 총손익을 계산한 다음 20종목의 총손익을 합친 금액이 가장 높은 유전자 개체를 택하는 방식으로 했다. 최적화 과정에서는 1종목당 1회 매수 금액을 1천만 원으로 고정해 테스트했지만, 현실 시장에서 보통의 투자자라면 늘어난 수익을 재투자할 것이다. 전체적으로 10% 수익이 발생했으면 다음 회차에서는 1천1백만 원어치를 매수하는 것이 일반적이다.

시뮬레이터에서 이처럼 누적 금액 방식을 적용하려면 매번 포트폴리오의 전체 손익을 같이 계산해야만 개별 종목의 진입 수량을 계산해 낼 수 있다. 하지만 포트폴리오 전체 손익을 계산하려면 20종목을 병렬로 처리해야 가능하고 연산이 복잡해져 처리 시간이 길어지는 단점이 있다. 또한, GA를 사용하는 것 자체가 어렵 치이기 때문에 정교하게 매개변수를 찾을 필요가 없다. 따라서 성능을 고려하여 1종목씩 처리한 다음 전체 결과를 하나로 합치는 방법을 택하였다.

2) 매매주문 시 발생하는 체결 오차 현상으로 원하는 가격에 매매할 수 없을 때 발생하는 비용, 현재가가 10,000원일 때 즉시 체결을 원하면 1 호가를 더 올려서 10,050원에 매수주문을 내야 체결 가능성이 커진다.

이렇게 고정 금액 방식으로 계산된 매매 내역은 시물레이션이 끝나면 원래 의도했던 대로 누적 금액 방식으로 재계산해서 평가하여야 선택의

<표 4> 고정 금액 방식으로 가공한 매매 내역

종목명	날짜	신호명	수량	가격(원)	거래금액 (수수료반영)	수익(원)	수익률
신화실업	15/01/02	B	829	12,050	-10,050,879	0	0.0
씨아이테크	15/01/02	B	15,625	640	-10,035,625	0	0.0
에이프로젠 KIC	15/01/02	B	5,449	1,835	-10,046,158	0	0.0
지코	15/01/02	B	21,598	463	-10,041,472	0	0.0
진양폴리	15/01/02	B	4,192	2,385	-10,038,876	0	0.0
삼화전자	15/01/02	B	2,803	3,567	-10,032,313	0	0.0
부산산업	15/01/02	B	470	21,250	-10,030,975	0	0.0
청호컴넷	15/01/02	B	3,305	3,025	-10,034,145	0	0.0
에쓰씨엔지니어링	15/01/02	B	12,239	817	-10,031,501	0	0.0
세하	15/01/02	B	26,455	378	-10,046,445	0	0.0
동양네트웍스	15/01/02	B	17,730	564	-10,037,449	0	0.0
체시스	15/01/02	B	13,736	728	-10,033,544	0	0.0
주연테크	15/01/02	B	23,474	426	-10,043,398	0	0.0
삼화전자	15/01/08	XL0	2,803	3,510	9,775,322	-256,990	-2.6
지코	15/01/13	XL0	21,598	470	10,078,707	37,235	0.4
국보	15/01/15	B	913	10,950	-10,062,995	0	0.0
동양네트웍스	15/01/15	XL0	17,730	661	11,643,202	1,605,753	16.1
....	....	....	....	....	....	....	....
체시스	18/01/02	XL1	10,090	1,010	10,089,496	60,217	0.6
아티스	18/01/02	XL1	3,710	2,700	9,948,365	-88,632	-0.9
에이리츠	18/01/02	XL1	1,827	4,915	8,925,671	-1,007,497	-11.1



오류를 줄일 수 있다. <표 4>는 종목별 매매 결과를 하나로 합친 후 날짜순으로 소팅한 결과 테이블이다. <표 4>의 매매 내역을 한 줄씩 읽어가면서 아래의 수식에 따라 수량 및 금액 등을 재계산하였다.

if signal == buy then

$$A_{\max} = \frac{Cash_{t-1}}{NoPos_{t-1}}$$

$$C = \frac{A_{\max}}{P_B \times (1 + F_B) + S_B}$$

$$A_B = C \times (P_B \times (1 + F_B) + S_B)$$

$$Cash_t = Cash_{t-1} - A_B$$

$$NoPos_t = NoPos_{t-1} - 1$$

if signal == sell then

$$A_S = C \times (P_S \times (1 - F_S) - S_S)$$

$$Cash_t = Cash_{t-1} + A_S$$

$$NoPos_t = NoPos_{t-1} + 1$$



$$G = Cash_t - Cash_0$$

$$Y = \frac{G}{Cash_0}$$

Where  $A_{\max}$  := 최대매수가능금액,  $C$  := 계약수

$Cash_0$  := 최초시점의 현금즉, 투입자금 규모

$Cash_t$  := 현재시점의 현금

$Cash_{t-1}$  := 직전매매시점의 현금

$NoPos_t$  := 현재시점의무포지션종목수

$NoPos_{t-1}$  := 직전매매시점의무포지션종목수

$A_B$  := 매수금액,  $A_S$  := 매도금액

$P_B :=$  매수가격,  $P_S :=$  매도가격

$F_B :=$  매수수수료율,  $F_S :=$  매도수수료율

$S_B :=$  매수슬리피지,  $S_S :=$  매도슬리피지

$G :=$  총손익,  $Y :=$  수익률

<표 5> 누적 금액 방식의 일별 잔고 수량 테이블

날짜	씨아이테크	나노메딕	국보	....	부산산업	남광토건	세하	진양폴리
15/01/02	15,570	11,940	0	....	468	0	26,337	4,176
15/01/05	15,570	11,940	0	....	468	0	26,337	4,176
15/01/06	15,570	11,940	0	....	468	0	26,337	4,176
15/01/07	15,570	11,940	0	....	468	0	26,337	4,176
15/01/08	15,570	11,940	0	....	468	0	26,337	4,176
15/01/09	15,570	11,940	0	....	468	0	26,337	4,176
15/01/12	15,570	11,940	0	....	468	0	26,337	4,176
15/01/15	15,570	11,940	928	....	468	0	26,337	4,176
15/01/16	15,570	11,940	928	....	468	0	26,337	4,176
....	....	....	....	....	....	....	....	....
17/12/27	33,038	0	0	....	0	3,567	17,671	15,342
17/12/28	33,038	0	0	....	0	3,567	17,671	15,342
18/01/02	0	0	0	....	0	0	0	0

고정 금액 방식으로 계산했을 때 최종수익률이 102.5%였는데 누적 금액 방식으로 재계산하니 151.0%가 나왔다<표 7>. 누적 금액 방식으로 계산하면 최대손실률<sup>3)</sup>은 커지고, 손익비<sup>4)</sup>와 평균손익비<sup>5)</sup>는 낮아진다. 보상

3) 누적 수익 곡선 상의 최고 수익 대비 최대자본인하액을 비율로 계산한 값이며 본 논문에서는 미실현손익을 포함해서 계산했다. 값이 적을수록 좋다.

4) 총이익/총손실, 1보다 커야 하며 클수록 좋다.

5) 평균 이익/평균 손실, 1보다 커야 하며 클수록 좋다. 승률과 매매방법에 따라 상대

비율<sup>6)</sup>은 경우에 따라 다르게 나올 수 있다. 최대손실금액 증가분보다 총 손익의 증가분이 크면 보상비율은 증가할 것이고 반대로 총손익의 증가분이 적으면 보상비율은 고정 금액 방식보다 줄어든 것이다.

<표 6> 종목별 진입 가격 테이블

날짜	씨아이테크	나노메딕	국보	....	부산산업	남광토건	세하	진양폴리
15/01/02	640	835	0	....	21,250	0	378	2,385
15/01/05	640	835	0	....	21,250	0	378	2,385
15/01/06	640	835	0	....	21,250	0	378	2,385
15/01/07	640	835	0	....	21,250	0	378	2,385
15/01/08	640	835	0	....	21,250	0	378	2,385
15/01/09	640	835	0	....	21,250	0	378	2,385
15/01/12	640	835	0	....	21,250	0	378	2,385
15/01/15	640	835	10,950	....	21,250	0	378	2,385
15/01/16	640	835	10,950	....	21,250	0	378	2,385
17/12/27	758	0	0	....	0	7,020	1,560	1,720
17/12/28	758	0	0	....	0	7,020	1,560	1,720
18/01/02	0	0	0	....	0	0	0	0

평가해야 한다.

6) 총손익/최대자본인하액, 1보다 커야 하고 클수록 좋다. 상대평가해야 한다.

<표 7> 누적 금액 방식으로 재투자했을 때의 일별 평가금액

(단위: 원, %)

날짜	씨아이테크	나노메딕	....	진양폴리	미보유	현금	평가손익	평가액	수익률
15/01/02	-100,894	-509,480	....	-423,112	6	60,010,907	-2,601,537	197,398,463	-1.3
15/01/05	54,028	-568,881	....	-506,215	5	50,011,905	-1,323,009	198,676,991	-0.7
15/01/06	-54,417	-723,325	....	-734,746	5	50,011,905	-445,200	199,554,800	-0.2
15/01/07	-193,846	-687,684	....	-1,191,810	5	50,011,905	-841,009	199,158,991	-0.4
15/01/08	208,949	-616,403	....	-942,502	7	71,804,851	1,301,832	201,301,832	0.7
15/01/09	224,442	-640,163	....	-921,727	7	71,804,851	2,553,042	202,553,042	1.3
15/01/12	115,997	-497,599	....	-921,727	7	71,804,851	3,912,644	203,912,644	2.0
15/01/13	54,028	-545,121	....	-963,278	8	81,842,026	3,771,870	203,771,870	1.9
15/01/14	-100,894	-865,889	....	-1,004,829	8	81,842,026	4,188,564	204,188,564	2.1
15/01/15	-85,401	-723,325	....	-1,067,156	8	83,214,877	5,371,083	205,371,083	2.7
....	....	....	....	....	....	....	....	....	....
17/12/27	13,009,663	18,660,977	....	-1,004,385	10	261,746,397	297,733,973	497,733,973	148.9
17/12/28	14,357,448	18,660,977	....	-470,100	10	261,746,397	300,287,195	500,287,195	150.1
18/01/02	14,751,922	18,660,977	....	-470,100	20	501,974,642	301,974,642	501,974,642	151.0

## 제 4 장 실험 결과

본 논문에서는 패턴을 이용한 전략의 유용성을 검증하기 위해 M&W 파동 패턴을 근간으로 20종목으로 구성된 포트폴리오를 운용했을 때의 성과 분석을 제시하였다. 전환점의 배열을 인지하기 위해 3가지 지그재그 방식으로 테스트했으며 매개변수 최적화는 GA를 이용하였고 매매전략의 견고함을 측정하기 위해 WFA를 사용하였다. 또한, 패턴 매매에 적합한

<표 8> 과거 테스트 수익률 비교

테스트명		시총 상위	시총 하위	변동성 상위	변동성 하위	베타 상위	베타 하위
단순보유		15.9	103.5	71.6	5.9	32.5	23.1
최소 변동률 지그재그	1	22.3	188.1	39.3	15.7	26.7	34.2
	2	23.9	166.0	70.6	18.7	25.9	47.5
	3	24.7	133.1	114.3	14.1	18.6	27.8
	4	17.5	198.3	65.2	18.2	43.8	43.4
	5	18.5	145.2	80.7	13.5	37.0	35.9
고저라인 지그재그	1	20.1	115.1	71.7	17.2	68.0	36.3
	2	22.8	153.0	104.8	23.2	67.0	39.0
	3	18.7	121.8	81.2	19.0	42.5	46.6
	4	22.6	133.9	80.9	19.8	41.2	34.7
	5	19.7	112.8	90.1	22.4	40.7	35.7
스윙 파동	1	15.1	125.6	67.6	15.5	57.8	28.9
	2	24.1	120.9	113.9	21.0	46.7	33.7
	3	19.4	134.2	129.9	19.2	56.9	37.5
	4	14.3	120.7	86.9	17.4	48.8	41.6
	5	21.4	107.9	130.2	24.8	33.8	42.7

종목의 특성을 찾기 위해 시가총액, 변동성 그리고, 베타 값 3가지를 기준으로 종목을 나누어 테스트하였다.

WFA가 반영된 결과를 측정하기 전에 2015년부터 2017년까지 GA에 의한 과거 테스트 결과와 보유 수익률을 비교함으로써 초과 수익이 발생하였는지를 보았다. 논리적으로 WFA의 수익률은 과거 테스트의 수익률보다 높게 나올 수 없다. 과최적화를 아무리 피하려 해도 과거 테스트에서 과최적화 요인을 완전히 제거할 수 없다. 그러므로 과최적화된 결과조차 보유 수익률을 웃돌지 않는다면 WFA를 진행할 필요가 없는 셈이다.

<표 8>은 6개 포트폴리오의 과거 테스트 결과이다. GA를 사용하여 최적화할 경우 난수를 이용하기 때문에 실행할 때마다 결과값이 달라진다는 단점이 있다. 그래서 포트폴리오별로 지그재그 방식 3개에 대해 5회씩 반복해서 실험해 보았다. 90개의 테스트에서 81개 테스트 결과가 단순 보유 수익률을 웃돌았다. 따라서 GA를 이용한 최적화 과정이 유효했다고 볼 수 있다.

<표 9>는 지그재그 유형별로 평균과 표준편차를 구한 값이다. 3가지

<표 9> 과거 테스트 수익률의 평균 및 표준편차

전환점 계산방식	시총 상위		시총 하위		변동성 상위		변동성 하위		베타 상위		베타 하위	
	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차
단순보유	15.9		103.5		71.6		5.9		32.5		23.1	
최소변동률 지그재그	21.4	2.9	166.2	24.7	74.0	24.4	16.0	2.1	30.4	8.9	37.8	7.0
고저라인 지그재그	20.8	1.6	127.3	14.8	85.8	11.2	20.3	2.2	51.9	12.8	38.5	4.3
스윙 파동	18.9	3.7	121.8	8.5	105.7	24.7	19.6	3.2	48.8	8.7	36.9	5.1

지그재그 방식 중 어느 것이 더 적합하였는지 확인하고자 했으나 매매 대상 포트폴리오별로 순위가 달라서 우열을 가리진 못 하였다.

<표 10>은 WFA까지 마친 전체 그룹의 시뮬레이션 결과를 수익률 기준으로 소팅(sorting)한 것이다. 음영 셀(cell)은 해당 항목에서 최적의 값을 표시한 것이다. 시가총액 하위 종목으로 구성된 2번 포트폴리오의 단순 보유 시 수익률이 다른 포트폴리오보다 높아 2번 포트폴리오의 결과들

<표 10> 시뮬레이션 결과

테스트명	총손익	수익률	총거래 횟수	승률	최대 손실률	손익비	보상 비율	평균 손익비	거래평균 수익률
p2w4z1	301,974,642	151.0	335	53.7	-13.5	1.7	4.6	1.5	6.1
p2w12z1	219,436,346	109.7	249	52.6	-10.8	1.6	4.7	1.4	6.8
p3w3z3	217,614,915	108.8	249	45.4	-18.1	1.7	3.0	2.0	6.6
p2w8z3	213,203,998	106.6	235	57.9	-14.3	1.7	3.4	1.2	6.7
p2w3z1	192,237,903	96.1	440	50.2	-19.5	1.4	2.6	1.4	3.4
p3w12z3	190,987,505	95.5	144	47.9	-22.4	1.8	2.3	1.9	11.6
p2w9z1	183,494,671	91.7	647	43.3	-23.8	1.3	1.8	1.7	2.3
p2w4z3	175,929,289	88.0	236	53.8	-12.6	1.8	4.3	1.5	5.9
p2w10z2	168,483,503	84.2	155	57.4	-17.6	1.8	2.2	1.3	8.8
p2w6z1	164,496,984	82.2	276	52.9	-14.7	1.5	3.0	1.4	5.0
p2w2z1	142,588,576	71.3	264	55.7	-16.6	1.4	2.4	1.2	4.4
p2w5z2	141,170,650	70.6	270	46.7	-19.0	1.5	2.0	1.7	4.5
p2w10z1	139,943,754	70.0	229	52.4	-20.0	1.4	1.9	1.3	5.5
p2w3z2	128,186,801	64.1	300	47.7	-17.9	1.4	1.9	1.6	3.9
p2w5z1	124,044,695	62.0	371	53.4	-13.4	1.4	2.7	1.2	2.8
....	....	....	....	....	....	....	....	....	....
p1w9z2	-34,203,459	-17.1	434	42.6	-20.5	0.7	-0.8	1.0	-0.8
p6w12z1	-36,667,003	-18.3	278	45.0	-33.4	0.7	-0.5	0.9	-1.3
p4w9z1	-48,505,450	-24.3	544	36.2	-29.6	0.6	-0.8	1.1	-1.0
p1w9z3	-50,317,013	-25.2	369	36.3	-27.3	0.5	-0.9	0.9	-1.6
p4w11z1	-54,311,761	-27.2	634	36.4	-34.3	0.6	-0.7	1.1	-1.0

이 상위에 올라와 있다.

전략의 우수성보다 종목 선정의 운이 좋았던 것으로 해석될 수 있으므로 결과에 대해 정보비율(information ratio, IR)<sup>7)</sup>을 비교해 보았다. 정보비율은 주로 펀드 평가에서 사용되며 초과 수익률은 펀드 수익률에서 벤치마크 수익률(benchmark rate)을 뺀 값이다. 본 실험 분석에서 벤치마크 수익률을 단순 보유 전략의 수익률로 보았다.

$$IR_p^M = \frac{\overline{R_p} - \overline{R_M}}{\sqrt{\frac{1}{n-1} \sum_{j=1}^n [(R_p^j - R_M^j) - (\overline{R_p} - \overline{R_M})]^2}}$$

Where  $n :=$  데이터(주간로그수익률)의 개수

$R_p^j :=$  포트폴리오의  $j$ 기의 주간로그수익률

$R_M^j :=$  단순 보유 전략의  $j$ 기의 주간로그수익률

$\overline{R_p} :=$  포트폴리오의 주간로그수익률의 평균

$\overline{R_M} :=$  단순 보유 전략의 주간로그수익률의 평균

수익률만 가지고 전략을 선택하면 실전 매매에서 재현되지 않아 낭패를 볼 수 있으므로 수익률을 제외한 다른 요소들로 평가하여야 한다. 이승용 외(2004)는 좋은 전략을 고르는 기준으로 승률 30% 이상, 손익비, 평균 손익비, 보상비율이 1보다 커야 한다고 하였다. 또한, 최대손실의 크기를

7) 펀드매니저의 능력을 측정할 수 있는 지표로 적극적인 투자 활동의 결과로 나타나는 초과 수익률과 적극적인 활동에 따른 수익률의 표준편차의 비율을 말하는데, Reward-to-Variability Ratio(RVR)라고 부르기도 함. 이 비율이 높을수록 더 좋은 투자 활동으로 판단한다.(출처:펀드닥터)



<표 11> 필터링된 최적화 결과

테스트명	수익률	승률	최대 손실률	보상비율	손익비	평균 손익비	CPC Index	정보비율	샤프비율
KOSPI	28.7	-	-15.8	1.6	-	-	-	-	0.7
p3w12z3	95.5	47.9	-22.4	2.3	1.8	1.9	1.7	0.3	1.1
p3w3z3	108.8	45.4	-18.1	3.0	1.7	2.0	1.5	0.4	1.2
p2w4z1	151.0	53.7	-13.5	4.6	1.7	1.5	1.3	0.4	1.5
p2w8z3	106.6	57.9	-14.3	3.4	1.7	1.2	1.2	0.0	1.3

확인해야 하는데 25% 손실이 발생하면 다음에는 33.33% 수익이 나와 원금을 회복할 수 있으므로 적어도 25% 이상 손실을 허용하면 안 된다고 하였다. 또한 Harris(2011)는 승률과 손익비, 평균 손익비 3개의 값을 곱한 값이 1.2보다 크면 좋다고 하였다.<sup>8)</sup>

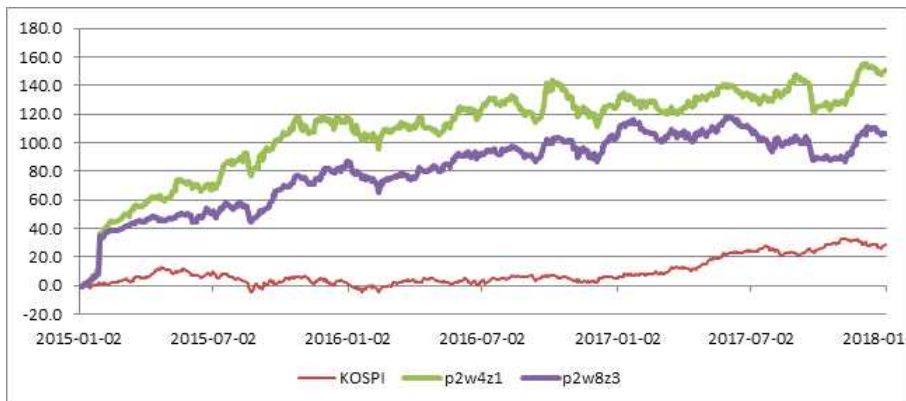
$$CPC = \text{손익비} \times \text{승률} \times \text{평균손익비}$$

$$\text{손익비} = \text{abs}\left(\frac{\text{총이익}}{\text{총손실}}\right)$$

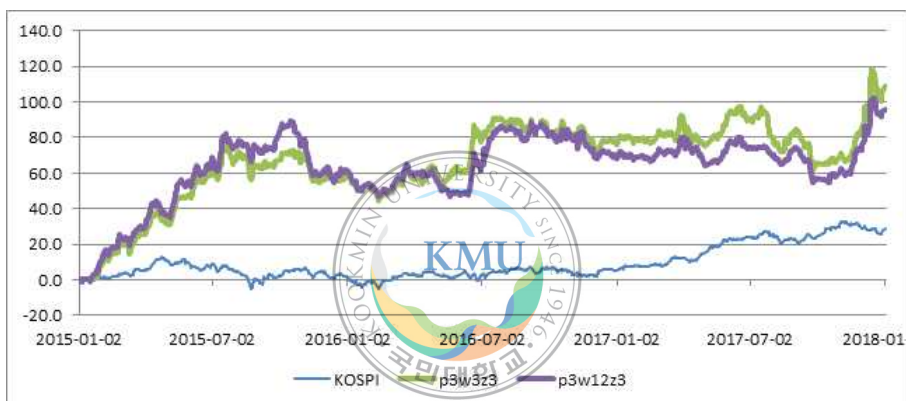
$$\text{평균손익비} = \frac{\frac{\text{총이익}}{\text{이익거래횟수}}}{\frac{\text{총손실}}{\text{손실거래횟수}}}$$

- 최대손실률의 절댓값 : < 25%
- 승률 : > 30%
- 손익비 : > 1
- 평균손익비 : > 1

8) CPC(Cardinal Profitability Constructs) Index



(a) 시가총액 하위 그룹



(b) 변동성 상위 그룹

<그림 21> 수익률 곡선 비교

- 보상비율 :  $> 1$
- CPC Index :  $> 1.2$
- 누적수익률 :  $>$  동기간 KOSPI 수익률
- 정보비율 :  $> 0$

<그림 21>는 4개 테스트 결과의 누적 수익 곡선을 나타낸 것인데, <표 11>에서 3번 포트폴리오(변동성 상위 종목)의 CPC Index 값이 2번 포트

폴리오(시가총액 하위 종목)보다 더 높았으나 수익 곡선은 2번 포트폴리오가 꾸준히 상승하고 있어 후자가 더 좋게 나왔다.

<그림 22>는 4개 포트폴리오의 보유 종목수 변동 추이를 나타내고 있



<그림 22> 보유 종목수 변동 추이

다. 종목수가 0이거나 특이한 구간이 없으므로 시뮬레이션 결과는 적절했다고 볼 수 있다.

샤프 비율은 수익 곡선의 진행 과정을 분석하는 지표이다.

$$S_p = \frac{\overline{R_p}}{\sigma_p}$$

Where  $\overline{R_p} :=$  포트폴리오 주간로그수익률의 평균

$\sigma_p :=$  포트폴리오 주간로그수익률의 표준편차

수익 곡선이 선형적으로 상승하였다면 이 수치가 높고 수익 곡선이 등락을 반복하면 표준편차가 커져, 이 수치가 낮다(이승용 외, 2004). 샤프 비율은 펀드 평가에서도 사용되는데 포트폴리오의 평균 수익률에서 무위험 자산의 평균 수익률을 빼 초과 수익률을 포트폴리오의 표준편차로 나눠서 구한다. 여기서는 수익 곡선의 선형성을 비교하기 위함이므로 무위험 자산의 평균 수익률은 고려치 않고 계산했으며 샤프 비율로 비교하면 시가총액 하위 종목 포트폴리오의 수익이 더 안정적이었다<표 11>.

## 제5장 결론 및 향후 연구

주가 패턴에 관한 다수의 기존 연구들은 주가 예측력이 높은 패턴을 찾아내는 데 목적을 두고 있다. 이렇게 의미 있는 패턴이 찾아지면 해당 패턴과 일치하는 시점을 또 찾는다. 그리고 해당 시점에 진입했다고 가정하고  $n$  일이 지났을 때의 성과를 측정한다. 이러한 접근법은 가상의 수익을 계산하는 것이기 때문에 현실과 많은 괴리가 발생할 수 있다.

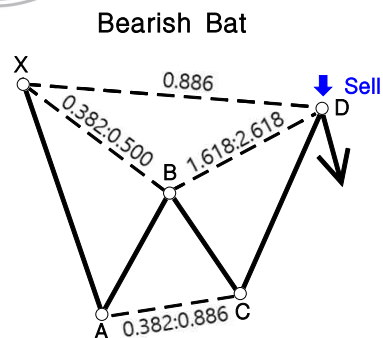
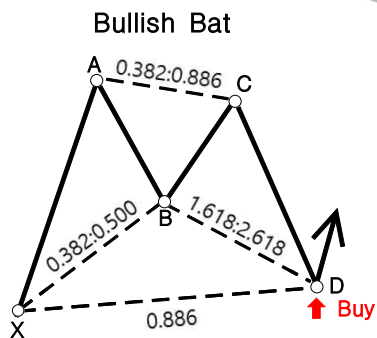
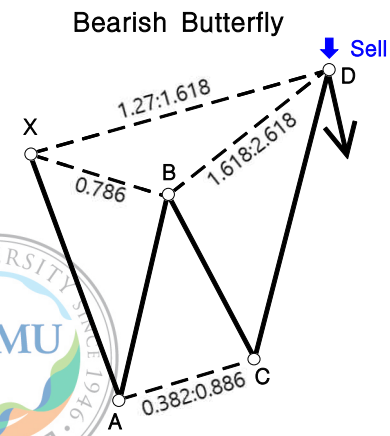
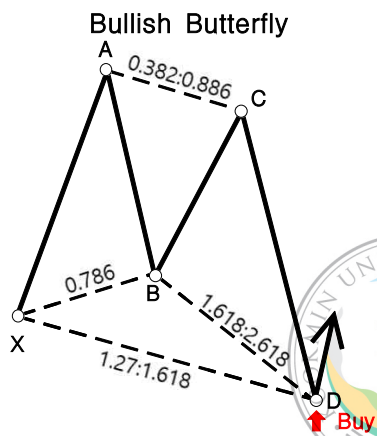
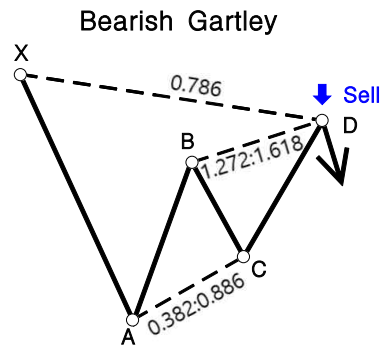
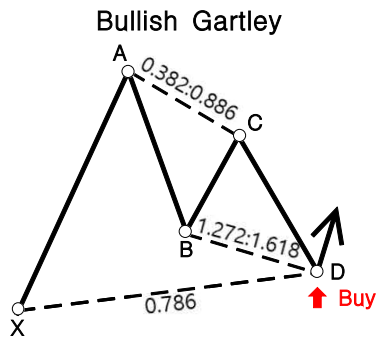
본 연구에서는 상승전환과 하락전환으로 정의된 32개 패턴을 10개 그룹으로 나누고 확률기반으로 선택하여 매매하는 방법을 실험하였다. 과거 성공률이 높았던 패턴이 미래에도 성공 확률이 높을 것이라 보고 해당 패턴이 발생했을 때 진입과 청산이 이루어지도록 한 것이다. 진입과 청산을 모두 반영해서 측정했기 때문에 실제 상황에 가깝다. Merrill의 M&W 파동 패턴 분류 방법을 사용한 이유는 5개의 전환점으로 거의 모든 패턴을 설명할 수 있는 단순함 때문이며 결과적으로 패턴을 정교하게 찾지 않아도 됨을 확인하였다. 성공률이 높은 패턴을 찾기 위해 현실적으로 전수 테스트가 불가능했으나 GA는 가장 적합한 해결 방법이었다. 또한, 테스트 구간과 적용 구간을 달리하는 WFA 기법을 적용하였기 때문에 시장 변화에 적절히 대응할 수 있었다.

6개의 종목군으로 나뉘서 실험한 결과에서 시가총액 하위 종목의 포트폴리오가 가장 성과가 좋았으며 변동성 상위 종목의 포트폴리오가 두 번째로 좋았다. 이는 차트상 패턴이 만들어지려면 어느 정도의 가격 움직임이 있어야 하지만 변동성이 클수록 더 좋은 것은 아니라는 것을 보여준다.

본 연구 결과의 한계는 가격 데이터만을 이용하여 매매가 100% 성립되었다고 가정하였기 때문에 실제로 거래에 적용했을 때 차이가 발생할 수 있는 점이다. 거래정지 또는 유동성 부족 등의 문제로 슬리피지보다 더 큰 비용이 발생할 수 있다. 그리고 실험에 소요되는 시간이 오래 걸리기 때문에 장기간의 데이터를 분석하지 못하는 것이 단점이다. 가격 데이터를 하나씩 읽어가면서 거래 내역을 만들어야 하는 실험 방법의 특성 때문에 소요 시간을 단축하는 것에 한계가 있다. 이번 실험에서는 120종목의 6년 치 개별 주가 데이터를 사용하였는데 해당 기간이 박스권 장세였기 때문에 상승 국면과 하락 국면에서도 KOSPI보다 월등한 수익이 나오리라 기대할 수는 없다. 그렇지만 WFA 기법을 적용하여 시장 국면이 바뀌면 매개변수들도 바뀌기 때문에 실험결과와 전혀 다른 결과가 나오지 않을 것이다.

향후 연구과제로는 본 연구의 방법을 개선해서 패턴 이론의 하나인 하모닉 패턴(harmonic pattern)<sup>9)</sup>으로 매매시스템을 만드는 것이다. 하모닉 패턴도 M&W 파동 패턴과 마찬가지로 5개 전환점으로 구성되어 있지만, 전환점과 전환점 사이의 반등, 반락 비율을 핵심 요소로 본다. 유의미한 패턴들이 이미 보고되어 있으나 M&W 파동 패턴보다 복잡하며 해석이 다양하다. 해외 상용 프로그램들도 구현 방법이 제각각이며 해외 시장 상황에 맞춰 발전되어 왔기 때문에 국내 시장과는 괴리가 있을 수 있다. 따라서 국내 시장 환경 맞는 값들을 찾는 것이 중요하다.

9) Gartley(1935)가 그의 저서 '주식시장의 이익'에서 처음 소개하였으며 이후 Carney를 비롯한 추종자들이 가틀리 패턴 외에 박쥐 패턴, 상어 패턴, 나비 패턴 등을 추가하면서 '하모닉 패턴'으로 발전하였다. 특징은 반등 또는 반락의 기준을 피보나치 비율로 설명하고 있다.



< 그림 23 > 하모닉 패턴 중 가틀리, 나비, 박쥐 패턴

## 참 고 문 헌

강성석, 이현준, 오경주. “유전자 알고리즘을 활용한 군집화 기반 펀드투자 전략.” 한국데이터정보과학회지, 제30권, 제2호, 2019, pp.285-297.

김명중. “유전자 알고리즘을 이용한 SVM 앙상블의 최적화 Optimization.” 한국경영학회 통합학술발표논문집, 2009, pp.1-14.

김선웅, 안현철. “Support Vector Machines와 유전자 알고리즘을 이용한 지능형 트레이딩 시스템 개발.” 지능정보연구, 제16권, 제1호, 2010, pp.71-92.

김원조. 『주식 실전 차트 분석』. 서울: 좋은땅, 2016.

김은충, 오경주. “은닉 마르코프 모델(HMM)과 유전자 알고리즘(GA)을 이용한 자산배분전략 제안.” 한국데이터정보과학회지, 제30권, 제1호, 2019, pp.33-44.

김정환. 『차트의기술』. 경기도:이레미디어, 2013.

김해균, 김성신. “퍼지시스템과 지식정보를 이용한 주가지수 예측.” 대한전기학회 학술대회 논문집, 2001, pp.2030-2032.

백진우, 조성준. “자기연상학습 신경망과 부호변수를 이용한 종합주가지수” 왼쪽 어깨” 패턴 검출.” 한국정보과학회 학술발표논문집, 제27권, 제2호(II), 2000, pp.320-322.

송진호, 최홍식, 김선웅. “기계학습을 활용한 상품자산 투자모델에 관한 연구.” 지능정보연구, 제23권, 제4호, 2017, pp.127-146.

신동하, 최광호, 김복창. “RNN과 LSTM을 이용한 주가 예측을 향상을 위한 딥러닝 모델.” 한국정보기술학회논문지, 제15권, 제10호, 2017, pp.9-16.

아사이 노보루. 『엑셀로 배우는 인공지능』. 우영운(역), 경기도: 제이펍, 2017.

유재필, 신현준, 김미희, 백종관. “기계학습과 데이터 시각화 기법을 이용한 주가 패턴 분석.” 정보화연구(구 정보기술아키텍처연구), 제14권, 제2호, 2017, pp.189-197.



윤재수. 『차트분석 무작정 따라하기』. 서울: 길벗, 2017.

이승용 외. 『7th 감각, 시스템 트레이딩 with CybosTrader』. 서울: 대신경제연구소, 2004.

이재영, 한치근. “패턴을 이용한 KOSPI200 예측 시스템.” 한국정보과학회 학술발표논문집, 제30권, 제2호( I ), 2003, pp.508-510.

이재원. “이동 평균선 패턴과 전환점 행렬에 기반한 주식 거래 시스템.” 정보과학회논문지 : 컴퓨팅의 실제 및 레터, 제18권, 제7호, 2012, pp.528-532.

이재원. “고변동 주가 패턴의 감독 학습에 기반한 주식 거래 시스템.” 정보과학회논문지 : 컴퓨팅의 실제 및 레터, 제19권, 제1호, 2013, pp.23-29.

이재원, 오장민. “인공지능 통합 다중 시뮬레이션에 의한 신경망 기반 주식 거래 시스템의 성능 최적화.” 정보처리학회논문지B, 제14권, 제2호, 2007, pp.127-134.

이종식, 안현철. “입력변수 및 학습사례 선정을 동시에 최적화하는 GA-MSVM 기반 주가지수 추세 예측 모형에 관한 연구.” 지능정보연구, 제23권, 제4호, 2017, pp.147-168.

이종우, 김유섭, 김성동, 이재원, 채진석. “패턴 매칭과 자동 규칙 생성에 기반한 2단계 주식 트레이딩 시스템.” 정보처리학회논문지B, 제10권, 제3호, 2003, pp.257-264.

이주상, 이상욱, 장석철, 안병하. “유전 알고리즘을 이용한 주식시장에서의 기술적 분석에 의한 매매에 관한 연구.” 대한산업공학회 추계학술대회 논문집, 2004, pp. 734-737.

정용관, 윤영섭. “인공신경망모형을 이용한 주가의 예측가능성에 관한 연구.” 재무관리연구, 제15권, 제2호, 1998, pp.369-399.

조지호. “주식투자 수익율의 예측에 관한 연구.” 한국증권학회지, 제11권, 제1호, 1989, pp.231-257.

조희연, 김영민. “유전자 알고리즘을 이용한 주식투자 수익률 향상에 관한 연구.” 정보시스템연구, 제12권, 제2호, 2003, pp.1-20.

최홍식, 김선웅, 박성철. “방향성매매를 위한 지능형 매매시스템의 투자성과분석.”  
지능정보연구, 제17권, 제3호, 2011, pp.187-201.

한혜론. 『패턴으로 찾아 지표로 매매하자』. 서울: 생활지혜사, 2000.

Bao, D. and Z. Yang, “Intelligent Stock Trading System by Turning Point Confirming and Probabilistic Reasoning,” *Expert Systems with Applications*, Vol.34, No.1(2008), pp.620-627.

Bulkowski, T. N., *Getting Started in Chart Patterns*, Wiley, New Jersey, 2006.

Cao, L. J. and F. E. H. Tay, “Support Vector Machine with Adaptive Parameters in Financial Time Series Forecasting,” *IEEE Transaction on Neural Networks*, Vol.14, No.6, 2003, pp.1506-1518.

Fama, E. F., “Efficient capital markets: A review of theory and empirical work,” *Journal of Finance*, Vol.25, No.2, 1970, pp.383-417.

Harris, S. J., *TradeStation made easy! : using EasyLanguage to build profits with the world's most popular trading software*, Wiley, New Jersey, 2011.

Hill, A., *Using the ZigZag Feature for Elliott Wave Counts*, StockCharts, 2010. Available at <http://stockcharts.com/articles/mailbag/2010/11/using-the-zigzag-feature-for-elliott-wave-counts.html> (Accessed 20 May, 2018).

Kamijo, K. and T. Tanigawa, “Stock Price Pattern Recognition-a Recurrent Neural Network Approach-,” *International Joint Conference on Neural Networks*, 1990, pp.215-221.

Leigh, W., N. Modani, R. Purvis and T. Roberts, “Stock Market Trading Rule Discovery Using Technical Charting Heuristics,” *Expert Systems with Applications*, Vol.23, No.2, 2002, pp.155-159.

Levy, R. A., “The Predictive Significance of Five-Point Chart Patterns,” *Journal of Business*, Vol.44, No.3, 1971, 316-323.

Liu, J. N. K. and R. W. M. Kwong, "Automatic Extraction and Identification of Chart Patterns towards Financial Forecast," *Applied Soft Computing Journal*, Vol.7, No.4, 2007, pp.1197-1208.

Merrill, A. A., "M&W WAVE PATTERNS," *Market Technicians Association (MTA) Journal*, 1980, pp.43-54.

Merrill, A. A., "M and W WAVES - MORE DATA." *Market Technicians Association (MTA) Journal*, 1984, pp.23-29.

Mitchell, M., *An Introduction to Genetic Algorithms*, The MIT Press, Massachusetts, 1996.

Nyemera, E. (n.d.). BACKGROUND ON M AND W PATTERNS. Retrieved from [emmanuelnyemera.com](http://emmanuelnyemera.com)

Pardo, R., *The Design, Testing, and Optimization of Trading Systems*, Wiley, New Jersey, 1992.

Pardo, R., *The Evaluation and Optimization of Trading Strategies*, Wiley, New Jersey, 2008.

Shiller, R. J., "The Use of Volatility Measures in Assessing Market Efficiency," *The Journal Of Finance*, Vol.36, No.2, 1981, pp.291-304.

SirForex. "MW PATTERNS IN FOREX: DOUBLE TOP AND DOUBLE BOTTOM." Available at <https://sirforex.net/blog/forex-mw-patterns-explained>. (Accessed 30 Apr. 2018.)

Wilder, J. W., *New Concepts in Technical Trading Systems*, Trend Research, North Carolina, 1978

Zbikowski, K., "Using Volume Weighted Support Vector Machines with Walk Forward Testing and Feature Selection for the Purpose of Creating Stock Trading Strategy," *Expert Systems with Applications*, Vol.42, No.4, 2015, pp.1797-1805.

## Abstract

### Development of a Stock Trading System Using Stock Price Patterns and Genetic Algorithms

*by Yang, Hoon Seok*

Graduate School of Business IT, Kookmin University,  
Seoul, Korea

Investors prefer to look for trading points based on the graph shown in the chart rather than complex analysis, such as corporate intrinsic value analysis and technical auxiliary index analysis. However, the pattern analysis technique is difficult and computerized less than the needs of users. In recent years, there have been many cases of studying stock price patterns using various machine learning techniques including neural networks in the field of artificial intelligence(AI). In particular, the development of IT technology has made it easier to analyze a huge number of chart data to find patterns that can predict stock prices. Although short-term forecasting power of prices has increased in terms of performance so far, long-term forecasting power is limited and is used in short-term trading rather than long-term investment. Other studies have focused on mechanically and accurately identifying patterns that were not recognized by past technology, but it can be vulnerable in practical areas because it is a separate matter whether the patterns found are suitable for trading. When they find a meaningful pattern, they find a point that matches the pattern. They then measure their performance after n days, assuming that they have bought at that point in time. Since this approach is to calculate virtual revenues, there can be many disparities with reality.

The existing research method tries to find a pattern with stock price prediction power, but this study proposes to define the patterns first and to

trade when the pattern with high success probability appears. The M & W wave pattern published by Merrill(1980) is simple because it distinguishes it by five turning points. Despite the report that some patterns have price predictability, there were no performance reports used in the actual market. The simplicity of a pattern consisting of five turning points has the advantage of reducing the cost of increasing pattern recognition accuracy. In this study, 16 patterns of up conversion and 16 patterns of down conversion are reclassified into ten groups so that they can be easily implemented by the system. Only one pattern with high success rate per group is selected for trading. Patterns that had a high probability of success in the past are likely to succeed in the future. So we trade when such a pattern occurs. It is a real situation because it is measured assuming that both the buy and sell have been executed.

We tested three ways to calculate the turning point. The first method, the minimum change rate zig-zag method, removes price movements below a certain percentage and calculates the vertex. In the second method, high-low line zig-zag, the high price that meets the  $n$ -day high price line is calculated at the peak price, and the low price that meets the  $n$ -day low price line is calculated at the valley price. In the third method, the swing wave method, the high price in the center higher than  $n$  high prices on the left and right is calculated as the peak price. If the central low price is lower than the  $n$  low price on the left and right, it is calculated as valley price. The swing wave method was superior to the other methods in the test results. It is interpreted that the transaction after checking the completion of the pattern is more effective than the transaction in the unfinished state of the pattern.

Genetic algorithms(GA) were the most suitable solution, although it was virtually impossible to find patterns with high success rates because the number of cases was too large in this simulation. We also performed the simulation using the Walk-forward Analysis(WFA) method, which tests the test section and the application section separately. So we were able to respond appropriately to market changes. In this study, we optimize the stock portfolio because there is a risk of over-optimized if we implement the variable optimality for each individual stock. Therefore, we selected the number of constituent stocks as 20 to increase the effect of diversified investment while avoiding optimization. We tested the KOSPI market by

dividing it into six categories. In the results, the portfolio of small cap stock was the most successful and the high vol stock portfolio was the second best. This shows that patterns need to have some price volatility in order for patterns to be shaped, but volatility is not the best.

**Key Words** : Stock Trading System, Stock Price Patterns, M&W Wave Patterns, Genetic Algorithm, Walk-Forward Analysis, Portfolio Optimization

