

Self-Organizing Network에서 기계학습 연구동향-I

Research Status of Machine Learning for Self-Organizing Network - I

권동승 (D.S. Kwon, dskwon@etri.re.kr)

지능형고밀집스물셀연구실 책임연구원

나지현 (J.H. Na, jhna@etri.re.kr)

지능형고밀집스물셀연구실 책임연구원/실장

ABSTRACT

In this study, a machine learning (ML) algorithm is analyzed and summarized as a self-organizing network (SON) realization technology that can minimize expert intervention in the planning, configuration, and optimization of mobile communication networks. First, the basic concept of the ML algorithm in which areas of the SON of this algorithm are applied, is briefly summarized. In addition, the requirements and performance metrics for ML are summarized from the SON perspective, and the ML algorithm that has hitherto been applied to an SON achieves a performance in terms of the SON performance metrics.

KEYWORDS SON, Machine Learning

1. 서론

이동통신망의 설치와 운영은 매우 복잡한 작업으로 계획, 설치, 테스트, 사전 실행 최적화, 실행 후 최적화, 성능 모니터링, 결함 완화와 결함 수정과 같은 많은 요소를 고려해야 한다. 이 모든 활동은 노동 집약적이며 비용이 많이 들고 오류가 발생하기 쉬우며, 약간의 간과로도 상당한 수익 손실을 야기하는 고객 불만이 발생할 수 있다. 따라서 사람의 개입을 최소화하면서 계획, 설치와 유지 관리

활동의 효율성 극대화를 위해 3GPP(Third Generation Partnership Project)에서는 LTE(Long Term Evolution)에서 자가 조직화 네트워크(SON: Self-Organizing Network)를 정의하고, SON 지원을 위해 표준에 기능, 인터페이스와 절차를 개발했다.

SON은 망 운영 작업에 사람의 개입을 줄이면서 망 용량, 커버리지와 서비스 품질을 최적화하여 운영과 자본 지출을 크게 줄이는 것이다. SON 아이디어는 망 계획, 구성과 최적화에 최소한의 수동 개입을 필요로 하는 통합된 자동 혹은 자율 프

* DOI: <https://doi.org/10.22648/ETRI.2020.J.350410>

* 이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임[No. 2020-0-009454, 5G 스물셀을 위한 인공지능 기반 자율구성 네트워크(SON 기술 개발)].



본 저작물은 공공누리 제4유형

출처표시+상업적이용금지+변경금지 조건에 따라 이용할 수 있습니다.

©2020 한국전자통신연구원

로세스로 통합하는 것이다. 따라서 비용 최적화를 위한 용량과 커버리지 최적화, QoS 개선, 그리고 에너지 효율 달성의 네 가지 목적 달성을 위해 SON은 자가 구성(Self-Configuration), 자가 최적화(Self-Optimization)와 자가 치유(Self-Healing)로 구성된다. 자가 구성은 기지국(BS: Base Station)의 설치·확장·업그레이드 중에도 필요한 기능이며, 노드 장애, 망 성능 저하 또는 서비스 유형 변경 같은 시스템 변경 시에도 필요하다. BS IP 주소, 인접 셀 리스트(NCL: Neighbor Cell List), 무선 액세스 파라미터 등이 포함된다. 자가 최적화는 초기 자가 구성 단계 이후 시스템의 효율적인 성능 보장을 위해 시스템 파라미터를 지속적으로 최적화하는 것으로 부하 밸런싱, 용량과 커버리지 최적화, 간섭 제어가 포함된다. 자가 치유는 망 장비의 고장 영향 완화를 위해 원격 탐지, 진단과 트리거링, 보상 또는 복구 조치를 포함하는 프로세스로서 자가 치유에 대한 일반적인 접근 방식은 모니터링, 진단 및 보상의 기본 요소로 구성된다.

사물인터넷, 사물 간 통신, 클라우드 컴퓨팅 등 발전으로 연결성과 트래픽 밀도의 기하급수적 증가에 대응하기 위해 5G는 4G 대비 대기 시간, 용량과 신뢰성과 같은 제한 사항 해결을 위해 Massive-MIMO(Multiple-Input Multiple-Output), 밀리미터 파, 새로운 물리계층 파형, 망 가상화, 제어 및 데이터 평면 분리, 망 밀집화 및 SON 기능 등이 포함되어 있다. 이 중 망 밀집화 개념은 최종 사용자에게 더 높은 데이터 속도와 낮은 대기 시간을 제공하지만, 관리할 셀 노드의 증가, 그 셀의 성능을 모니터링으로 수집된 데이터 증가로 운영자는 초밀집 셀 간 조정, 구성과 관리에서 어려움이 발생한다. 따라서 망에 탄력성 제공, 망의 전반적인 복잡성, CAPEX와 OPEX 감소 그리고 망의 조정, 최적화와 구성 절차 단순화에 망 내부에 지

능 제공이 필요하다.

본 고의 II 장에서는 3G부터 SON에서 기계학습(ML: Machine Learning)에 대한 초기 연구 현황을, III 장에서는 ML을 SON에 적용된 ML 기술을 소개하고, IV 장에서는 SON에 대한 ML 요구사항을 서술하였다.

II. SON에서 Pre-Machine Learning 연구 현황

R. Barco 등은 무선액세스망에서 진단 전문가에 의해 수동적으로 진행되는 문제 검출, 원인 식별과 문제 해결 프로세스의 자동 진단·식별·해결을 하기 위해 Bayesian 분류기 기반으로 제안했다[1]. 알람과 망 성능 지표를 고려한 GSM·GPRS 망에서 문제의 자동 진단을 위한 추론 엔진과 지식 기반 모델을 연구했다.

이동통신망 운영은 실제 측정된 서비스 성능 분석 결과값을 해당 서비스 핵심성과지표(KPI: Key Performance Indicator)의 품질 기준과 임계값과 비교하여 망 구성을 수정하는 반복적인 품질 프로세스라 할 수 있다. J. Laiho 등은 능동적이고 성공적인 동작을 위해 효과적 분석 방법이 선결 조건이라고 판단하고, 이 프로세스 중 분석과 시각화를 위해 다중 KPI를 동시에 처리하는 자가 조직화 맵(SOM: Self-Organizing Map)을 W-CDMA에 적용했다[2].

O. G. Aliu 등은 비용 효율적인 방식으로 사용자와 운영자의 요구사항 충족을 위해 4G에 지능과 자율적 적응성의 추가가 필요하다는 근거로 다음을 제시했다[3]. 첫째, 사용자 이동성과 다양한 무선채널로 이동통신시스템은 시간과 장소에 따라서 무선 자원의 낮은 이용 혹은 혼잡과 열악한 QoS를 초래하는 무선 자원의 과잉 이용이라는 어

려움이 있다. 둘째, 플러그 앤 플레이로 설계된 펌토셀은 인접 매크로셀의 성능을 저하시키는 심각한 간섭을 일으킬 수 있다. 셋째, 미래의 거대 규모의 이동통신을 고려하면, 장비의 수명주기 동안 필요한 주기적 수동 최적화 방식은 비효율적일 것이고, 또한 시스템 복잡성이 증가하면 인적 오류가 커져서 복구와 복원 시간이 길어진다. 넷째, 성능 향상 외에도 시스템의 구성, 시운전, 최적화, 유지 보수, 문제 해결과 복구에 필요한 고가의 숙련된 인력을 최소화할 필요가 있다.

S. Bi 등은 4G 상용화 후 급증하는 모바일 데이터 트래픽에 근본적 대처를 위해 빅 데이터 시대에 걸맞은 성능 향상을 위해 빅 데이터 트래픽을 효율적으로 처리하는 빅 데이터 인지 무선통신망 개념을 제안하고, 빅 데이터 해석의 핵심 톨로 ML을 제안했다[4].

M. Peng 등은 3GPP LTE-Advanced의 이기종 망(HetNets: Heterogeneous Networks) 표준 실현에서 전체 관리 비용 최소화과 전체 운용 효율 향상을 위해 동일 주파수 간섭 관리 복잡도와 운용 비용을 줄이는 데 SON 기술이 필수적임을 인지하고, 모니터링, 분석, 계획, 평가, 실행과 지식화로 구성된 자동화 관리 기반으로 자가 구성과 자가 최적화 SON 구조를 제안 평가했다[5]. 자가 구성에서 셀 식별자와 무선 자원 관리를 대상으로 연구한 결과 여전히 IP 주소, 인접 셀 리스트와 다중 안테나 모드에 대한 고도화된 자가 구성 체계 설계는 여전히 어려운 문제라고 했다. HetNets에서 부하 밸런싱과 간섭 제어의 자가 최적화에서 오버레이와 언더레이 HetNets의 용량과 커버리지는 이론적 분석과 시뮬레이션을 기반으로 향상시켜야 된다고 했다. 그리고 HetNets의 SON에서 자가 검출과 자가 치유도 중요하지만, 이기종 통신 개체를 통해 어떻게 outage 감지하고 성능 손실을 보상하는

가의 문제도 지적했다.

R. Barco 등은 결합 검출과 진단 그리고 outage 보상메커니즘을 포함한 자가 치유에 대한 연구가 상대적으로 부족한 이유로 두 가지를 지적했다[6]. 첫째, 이동통신망의 일반적인 결합과 그들의 주요 증상을 정의하기 어려워서 문서화되어 있지 않고 소수의 전문가에 의존하는 점, 둘째, 대부분의 결합 모델링이 너무 복잡하거나 실제 동작하는 망에서 결합 기록을 얻기가 어려운 점을 들었다. 그리고 이동통신망에서 자가 치유를 위한 통합된 프레임워크로 정보 수집, 결합 검출, 결합 식별과 진단 그리고 결합 복구를 제안했고, 각 단계 별로 도전 이슈도 언급했다. 첫째, 자가 치유 기능 실현기술로는 진단을 위한 Bayesian 기술, 결합 식별을 위한 통계적 학습 등 ML에 대한 연구 필요성을 언급했다. 특히 진단 시에는 연속적인 KPI를 어떻게 모델링하고, 과거 데이터로부터 어떻게 모델을 학습할 것인가, 그리고 결합 복구 시에는 인접 셀 파라미터를 어떻게 자동 조절할 것인가 등 이슈를 제기했다. 둘째, HetNets에서 검출과 진단이 매크로셀과 펌토셀에서 모두 고려해야 하는데 펌토셀에서 문제가 매크로셀에 의한 경우 등 문제를 제기했다. 셋째, 유럽의 SOCRATES 프로젝트는 자가 최적화 기능 간 조정을 연구했으나 실제 망에서는 임의의 케이스가 동일한 자가 최적화 모듈 내의 기능 간만 아니라 다른 모듈 간에도 상호 연결되어 있었다. 즉 일부 주요 파라미터는 자가 구성, 자가 최적화와 자가 치유의 영향을 모두 받는다. 즉, 자가 치유에는 다른 SON 프로세스(자가 최적화와 자가 구성) 간 조정 연구가 필요하다. 따라서 자가 치유는 해당 논문의 참조 모델로 독립적으로 연구되었지만, 다른 SON 모듈, 특히 자가 최적화와 함께 연구되어야 함을 언급했다. 셀에 문제를 일으킬 수 있는 결합 원인 중 하나는 잘

못된 무선 파라미터인데, 자가 최적화가 완벽하면 잘못된 파라미터에 문제가 발생하지 않고, 자가 치유 프로세스에서 잘못된 파라미터로 인한 결합이 확인되면 해당 파라미터에 영향을 미치는 자가 최적화 알고리즘과 함께 조정해야 한다. 넷째, 통상 결합 검출과 진단을 위한 주요 입력 데이터는 운용유지로부터 얻은 알람과 KPI인데, BS와 단말의 온라인 측정과 현장시험에서 얻은 상황 정보와 데이터가 있으므로, 진단 정확도 향상을 위해 다른 소스와 다른 시간대의 정보를 어떻게 결합할 것인가가 이슈다. 다섯째, 자가 치유 설계 시 가장 어려운 점은 기존 망 시뮬레이터는 적합하지 않고 실제 망에서 시험하는 것은 시간이 많이 소요되므로 설계한 알고리즘을 어떻게 평가할 것인가이다. 구체적으로는 결합 원인과 KPI 간 관계를 통계적 모델링을 하거나 결합 모델링에 간단한 가정을 적용하는 것으로 자가 치유 프로세스 성능 평가 방법을 설계하는 것이다.

SON 기능들은 종종 독립적 기능으로 설계되지만, 동시에 동작될 때에는 그 기능들 간 상호작용은 항상 예측할 수 있는 것은 아니다. 예를 들어, 자가 최적화 메커니즘은 관리 혹은 제어 평면에 구현되는 제어 루프로 볼 수 있으며, 관리 평면에 구현되면 중앙집중 SON, 제어 평면에 구현되면 분산 SON이다. 그런데 다중 제어 루프가 설치되면 충돌, 안정성과 성능 문제가 발생하므로 충돌 해결, 조정과 다중 SON 기능 관리 프레임워크 이슈는 점점 더 중요해진다. 중앙집중 SON에서 조정 문제는 다목적 최적화로 취급하여 관리평면에서 실현하는 것으로 연구되었으나, 분산 SON에서 조정 문제에 대한 연구는 적었다. 따라서 A. Tall 등은 제어 평면에서 실제로 구현 가능한 SON 조정 메커니즘으로 완전 분산 조정인 경우 조정 문제를 제어 이론과 통계 근사 기반 프레임워크를

사용하여 Convex 최적화 문제로 수식화하고 분석했다[7]. 제안 메커니즘은 많은 수의 제어 루프를 처리하면서 안정성을 강화하는 기능이라 주장했다.

I. Karla[8] 등은 4G에서 다양한 셀과 망 구성을 위한 많은 파라미터 관리와 최적화 과정은 많은 수의 파라미터가 연결되어 있어서, 서로 다른 상호작용 메커니즘이 서로 영향을 미치므로 이 복잡한 구성과 최적화 문제 관리는 특정 유스 케이스에만 적용되도록 개발된 기존 SON 개념만으로 부족하다고 했다. 해결책으로 이론적 평가와 수학적 문제 공식화와 반복 학습을 통한 KPI의 지속 최적화를 위한 고전적인 학습 기법을 언급했다. 간단한 유스 케이스로는 피드백 루프, 인지 주기 및/또는 강화 학습을 기반 SON 기술을 사용하는 데, 동시에 최적화되어야 하는 서로 다른 어느 정도 연결된 파라미터의 양에 따라 거대한 후보 파라미터 공간이 초래되므로, 피드백 측정을 위해 실행 시스템에서 시험할 파라미터 구성의 선택사항을 줄이기 위해 Genetic 프로그래밍과 같은 지능형 검색 알고리즘 사용을 제안했다. 그러나 비교적 짧은 시간 단위로 해결될 필요가 있는 복잡하고 고도로 연결된 SON 유스 케이스에서는 신속하게 작동하는 SON 방식이 필요하다. 실행 시스템상에서 파라미터 구성을 시험 없이 잠재적 후보 파라미터의 미래 성능을 적어도 상당히 높은 확률로 예측할 수 있도록 예상되는 시스템 효과에 대한 명확한 이해를 기반이 요구되는 고도화 방법이 필요하다. 주요 도전은 각 개별 셀에 대해 매우 다양하고 무계획적이고 상호작용하는 환경을 이해하는 그런 시스템을 만드는 것이다. 즉 고도로 연결된 SON 환경에서 반응 속도, 확장성과 강건성을 보장하기 위해 사전 시험 없이 양호하고 안정적인 파라미터 구성을 충분히 높은 확

물로 예측할 수 있는 시스템으로 분산 SON 방식을 제안하였다. 기존 SON에 자체 학습 기능을 추가한 분산 SON 방식은 직접적인 피드백 없이 오프라인 계산을 사용하는 자체 학습 예측 모델을 통해 고도로 상호작용하는 셀 구성 파라미터를 구성하고 최적화한다고 했다.

자가 조정은 하나의 자가 조직화(SO: Self-Organizing) 기능 동작이 원래 의도한 목적이 다른 SO 기능에 영향을 주는 상황에서 필수이다. 따라서 3GPP Rel. 11에서 “LTE SON 조정 관리” 표준화 작업을 하였다. 한편 당시 망 노드 변이, 성능 메트릭, 셀 outage, 측정과 간섭 관점에서 SO 기능 충돌과 조정의 분류를 위한 포괄적인 프레임워크가 제안된 적이 없었다. H. Y. Lateef 등은 망 토폴로지 변이, KPIs, 출력 파라미터의 방향과 크기, 측정과 논리적 종속성을 기반으로 SO 기능 충돌을 식별하고 주석을 달고 분류하였으며, 자가 조정의 가능한 진화로 MRO(Mobility Robustness Optimization), MLB(Mobility Load Balancing) 기능의 TCA(Trigger-Condition-Action) 정책을 제안했다 [9]. 또한 자가 조정과 자가 최적화 간 최적 상호작용법을 식별했다.

P. Waino 등은 거리 수준의 실외 혹은 그 규모의 실내에 5G 소형셀 배치 시 기존 트랜스포트 망(백홀)의 효과적·경제적 이용이 사업자들의 주요 과제를 인식하고, 5G에서도 소형셀이 서로를 인식하고 망 토폴로지를 자체 구축하는 자가 구성 방법으로 5G 실외 소형셀 배치에 최적화된 새로운 무선백홀 솔루션을 제안했다[10]. 이 망 토폴로지의 핵심은 지능 엔진인데, 이 엔진은 세부적인 망 링크 계획과 무선백홀 망과 관련된 설계 작업을 최소화하면서 무선백홀 망 구축과 전송 리소스 설정 단계를 자동화하는 것이다.

III. SON에 적용된 ML 기술

SON에 적용된 ML 기술로는 지도 학습, 비지도 학습, 컨트roller, 강화학습(RL: Reinforcement Learning), Markov Models, Heuristics, Dimension Reduction, Transfer Learning(TL)이 있다. 여기서는 각 알고리즘의 개념과 SON의 어느 범주에 적용되었는지 정리해 보았다.

1. 지도 학습

지도 학습은 입력-출력 관계에 기초로 데이터에 대한 모델을 결정한 후, 새로운 입력 데이터 세트를 그 학습된 모델에 입력하여 출력을 예측하는 것이다.

Bayes 분류기는 가설의 확률이 새로운 증거에 비추어 어떻게 영향을 받는지 이해하기 위한 조건부 확률을 계산하기 위한 확률과 통계 분석인 Bayes 이론을 바탕으로 한 간단한 확률적 분류기이고, 입력이 서로 독립적이라는 가정에서 동작하는 것을 Naive Bayes 분류기라 한다. SON에서 결합 검출과 결합 분류에 이 Bayes 분류기 개념을 적용 연구되었다.

k-Nearest Neighbor(k-NN) 알고리즘은 관측값과 결과의 밀바탕에 있는 결합 분포를 알 수 없는 문제에 적용되는 것으로 비분류된 샘플을 가진 특정 클래스에 얼마나 많은 이웃이 있는가에 기반으로 새 샘플을 분류하는 것이다. k-NN은 통상 outage 혹은 sleeping 셀을 검출하는 SON의 자가 치유 영역에 적용 연구되었다.

신경망(NN: Neural Networks)은 인간의 뇌에서 보이는 것과 동일한 행동을 컴퓨터로 모사하는 시도로서 다중 계층 퍼셉트론(MLP: Multi-Layer

Perceptron)이라고도 부른다. 인간의 뇌는 매우 복잡한 비선형 및 병렬 계산을 항상 수행하는 복잡한 기계, 즉 NN에서 뉴런의 동등한 구성 요소인 노드 활성화 기능을 사용하여 비선형 계산을 수행하고, 가변 링크 가중치로 서로 연결되어 인간의 뇌에서 뉴런이 연결되는 방식을 모사한다. 노드 수와 히든 레이어 수를 변경하여 NN은 매우 복잡한 기능을 매핑하고 매우 우수한 성능을 달성할 수 있다. SON에서 NN은 대부분 분류기로 사용했고, 자가 최적화와 자가 치유에 적용 연구되었다.

Support Vector Machine(SVM)은 다른 클래스 간 거리 최대화를 위해 선형 또는 비선형 매핑을 이용하여 입력 세트를 보다 높은 차원의 특징 공간으로 매핑하는 것이다. SVM이 다른 클래스 간에 가장 큰 마진을 생성하는 초평면을 찾는 것이므로 SVM은 큰 마진 분류기로도 알려져 있다. SON에서 SVM은 자가 최적화와 자가 치유에 적용 연구되었다.

Decision Trees(DT)는 원 데이터의 부분 집합을 그 아래 부분 집합으로 반복적 분할로 구성하는 것으로 원래 데이터 기반으로 최상의 분류(분류 문제) 또는 값(회귀 문제)을 결정할 수 있도록 분할된 세트로 만든다. 분할이라는 것은 하위 분할에 포함된 데이터가 상위 노드의 데이터보다 더 순결하게 각 분할을 선택하는 것이다. SON에서 DT는 기본적으로 자가 최적화와 자가 치유에 적용 연구되었다.

추천 시스템으로 Collaborative Filtering(CF)는 다른 사용자 의견에 따라 사용자에게 제안하는 알고리즘으로 다른 사용자의 데이터베이스(DB: Data Base)를 기반으로 현재 사용자의 아이템 세트를 예측하는 것이다. SON에서는 자가 치유에 적용 연구되었다.

2. 비지도 학습

비지도 학습은 입력 세트만으로 각 관측에 대한 정답 혹은 오류 정도를 제공하는, 감독자 없이 출력을 올바르게 추론하는 것이다. 이 학습에는 레이블이 없는 입력 데이터 세트만으로 결과를 정확하게 학습해야 한다. 비지도 학습은 SON에서 운영 파라미터 구성, 자원 최적화, HO 관리, 이동성, 부하 밸런싱, 결합 검출, 셀 outage 관리, sleeping 셀 관리에 적용 연구되었다.

K-Means는 레이블이 없는 데이터 세트의 중심과 클러스터를 찾는 데 매우 유용한 클러스터링 알고리즘으로 초기 데이터 세트와 원하는 클러스터 수의 두 파라미터만 필요하다. 이 알고리즘 설치가 매우 쉽고 빠른 장점으로 인기가 있다. SON에서 K-means는 이동성 최적화, 자원 최적화, 결합 검출과 셀 outage 관리에 적용 연구되었다.

SOM 기술은 어떤 차원의 입력 신호라도 1~2차원의 이산 맵으로 변환하는 것을 목표로 일련의 데이터 항목에서 유사성 관계를 시각화한다. SOM 기본 속성으로 차원 축소 기술로 볼 수도 있고, SOM은 높은 차원의 데이터를 낮은 차원으로 순서대로 매핑하므로 원래 데이터에서 나타나는 복잡하고 비선형 관계를 낮은 평면에서 간단한 기하학적 관계로 변환시킨다. SOM은 SON에서 운영 파라미터 구성, 커버리지와 용량 최적화, HO 관리, 자원 최적화, 결합 검출과 셀 outage 관리에 적용 연구되었다.

비정상 검출기는 데이터에서 관찰된 특정 패턴과 일치하지 않는 데이터 포인트를 식별하는 것으로, 이 포인트를 비정상이라고 하며 일반적으로 시스템의 일반적인 동작과 적어도 다르거나 잘못된 것을 의미한다. 지도 학습, 세미 지도 학습과 비지도 학습이 가능하지만, SON에서 가장 일반적인

유형이 비지도 학습이다. 일부 알고리즘은 초기 데이터의 통계 측정과 새로운 데이터 포인트가 초기 분포로부터 얼마나 멀리 있는가에 의존한다. 다른 기술은 포인트 세트를 둘러싼 밀도와 이 영역이 얼마나 조밀한지에 따라 새로운 포인트가 정상 또는 비정상인지 표시된다. 또 다른 알고리즘은 새로운 포인트와 훈련된 데이터 간 상관 측정 혹은 단순한 규칙 세트로부터 편차에 의존한다. SON에서 비정상 검출 알고리즘은 주로 자가 치유에 사용되어 비정상적인 망 동작, 결합 분류를 검출하고 셀 outage 관리를 수행한다.

3. 컨트롤러

피드백 컨트롤러는 입력과 출력 간 규정된 관계 유지를 위해 지속적으로 파라미터를 조정하는 입력-출력 간 피드백 메커니즘이다.

컨트롤러는 지능형 알고리즘에 속하지 않지만 단순성과 구현 용이성으로 SON에 광범위하게 사용되었다. 즉 SON의 모든 도메인인 NCL 구성, 무선 파라미터 구성, 커버리지와 용량 최적화, HO 최적화, 부하 밸런싱, 자원 최적화, SON 기능 조정, 결합 검출과 셀 outage 관리에 적용 연구되었다. 이 컨트롤러는 현 입력-출력 기능과 원하는 기능 간 측정된 예러 기반으로 파라미터를 변경하므로 보다 정교하고 지능적인 방법을 적용한 다른 기술만큼 강력하지는 않다.

Fuzzy 로직 컨트롤러는 부분 진실을 나타내는 로직인 Fuzzy 로직을 사용하는 것으로, 즉 두 개의 이진 논리(0과 1) 사이에 보간을 적용하는 것이다. 이 컨트롤러는 표준 이진 로직 컨트롤러보다 granularity가 우수하므로 일반적으로 피드백 컨트롤러보다 FLC를 통해 보다 상세하고 복잡한 솔루션을 얻을 수 있다. 일반적인 Fuzzy 컨트롤러에

는 시스템의 현재 입력을 Fuzzy 로직 언어로 변환하는 Fuzzifier, 시스템의 입력과 출력 간 매핑을 정의하는 규칙 집합을 적용하는 추론 엔진과 모든 규칙을 집계하여 수량화 가능한 결과를 생성하는 Defuzzifier로 구성된다. Fuzzy 컨트롤러는 SON에서 HO 최적화, 부하 밸런싱과 같은 자가 최적화와 자가 치유 문제, 자원 최적화와 결합 검출에 적용 연구되었다.

4. Reinforcement Learning

RL은 주변 환경과 상호작용하고 현재 상태와 주변 환경 상태를 감지한 후, 행동을 선택하는 에이전트라는 개념을 기반으로 한다. 행동과 그 결과에 따라, 에이전트는 취한 행위가 좋으면 보상을 받고, 행위가 나쁘면 페널티를 받는다. 일반적으로 RL 시스템은 정책, 보상기능, 가치 함수, 환경 모델의 4단계로 나뉜다.

- 1) 에이전트가 취한 행위로 상태 매핑을 책임지는 정책
- 2) 현 상태를 평가하고 이전에 취해진 행위 결과에 따라 보상 혹은 페널티를 주는 보상 기능
- 3) 에이전트가 장기적으로 상태를 평가할 수 있는 가능성으로 미래에 선택된 상태로부터 예상되는 보상을 평가하는 가치 함수
- 4) 에이전트가 수행할 수 있는 상태와 가능한 행동을 결정하는 환경 모델

RL 보상 메커니즘 때문에 에이전트가 시스템에서 다른 조치를 취해서 결과를 탐색하는 것이 더 좋은지 혹은 현 지식을 유지하고 현재 알려진 행위의 보상을 극대화하는 것이 더 좋은지 중 결정해야 할 트레이드 오프가 있다.

가장 많이 사용되는 RL 알고리즘으로 시스

템이 가장 좋은 정책을 찾기 위해 Q 함수를 사용하는 Q-Learning(QL)과 Fuzzy Logic과 결합된 Q-Learning, 즉 Fuzzy Q-Learning(FQL)이 있다. SON에서 RL은 무선 파라미터 구성, 커버리지와 용량 최적화, HO 파라미터 최적화, 부하 밸런싱, 자원 최적화와 셀 outage 관리에 적용 연구되었다.

5. Markov Models

이 통계적 모델은 Markov 속성을 준수하면서 주로 랜덤하게 변경되는 시스템에 사용되며, 미래 상태의 조건부 확률 분포는 현재 상태값에만 의존하며 이전의 모든 값과는 독립적이다. 몇 가지 다른 Markov 모델 중 SON에서 가장 일반적으로 적용되는 모델은 Markov Chains(MC)과 Hidden Markov Models(HMM)이다. MC와 HMM 주요 차이점은 시스템 상태의 관찰 가능성으로 상태가 완전히 가시적이면 MC가 최선의 선택이며, 상태가 일부만 가시적이거나 전혀 표시되지 않으면 HMM을 선호한다.

Markov 모델은 SON에서 주로 자기 최적화 및 자가 치유에 적용 연구되었는데, 이동성 관리, 자원 최적화, 결합 검출과 셀 outage 관리이다.

6. Heuristics

휴리스틱 알고리즘은 기본적으로 주어진 시간에 시스템에 대한 최선의 결정을 하기 위해 특정 지침이나 규칙을 따르는 간단한 알고리즘으로 일반적으로 특정 문제에 대한 알려진 솔루션이 없거나 솔루션 계산에 비용이 너무 많이 드는 경우에 적용된다. 휴리스틱을 사용하면 근사적이고 차선의 솔루션을 찾을 수 있다. 간단한 예는 문제에 대한 솔루션을 계산할 수 없는 경우에는 brute-force

검색을 사용한다. 휴리스틱의 다른 종류인 메타 휴리스틱은 휴리스틱과 마찬가지로 일련의 기본 규칙을 따르지만 이전의 접근 방식과 달리, 보다 복잡하고 더 수준이 높아서 단순 휴리스틱보다 더 최적화된 솔루션을 얻을 수 있다. SON에서 휴리스틱은 커버리지와 용량 최적화, 부하 밸런싱에서 주로 자기 최적화에 적용 연구되었다.

휴리스틱의 또 다른 유형은 진화와 자연 선택과 같은 자연의 개념으로부터 영감을 얻은 Genetic Algorithms(GA)인데, 솔루션 군을 진화시키고 몇 세대 이후 최고의 솔루션을 찾기 위해 진화와 적자생존 메커니즘을 사용한다. GA는 매우 간단하지만 복잡한 문제에 대한 해결책을 찾을 수 있을 뿐만 아니라 비 결정적 문제도 해결할 수 있다. SON에서 이 알고리즘은 무선 파라미터 구성, 커버리지와 용량 최적화, HO 최적화, 부하 밸런싱, 자원 최적화, 셀 outage 관리 등 모든 종류의 문제 해결에 적용 연구되었다.

7. Dimension Reduction

차원 축소는 특징 선택 혹은 특징 추출의 두 가지 형태가 있다. 특징 선택은 초기 특징 세트에서 최상의 혹은 가장 유용한 특징만 선택하는 알고리즘이고, 특징 추출은 보다 유용하고 덜 중복되는 속성을 생성하기 위해 초기 특징 세트에 적용된 변환에 의존한다. 이 차원 축소 기술의 주요 동기는 분류기의 복잡성을 줄이고, 알고리즘 성능을 향상시키고 초기 데이터 세트에서 중복성과 유용성이 떨어지는 데이터를 제거해서 더 좋은 일반화 제공에 사용된다.

SON에서 가장 많이 사용되는 차원 축소 기술은 Principal Component Analysis(PCA), MCA(Minor Component Analysis), Diffusion Maps(DM),

Multi-Dimensional Scaling(MDS)이다. 이 모든 기술은 원 데이터 세트에 특정 유형의 변환을 적용하여 다른 공간으로 변환한다. PCA와 MCA는 변환된 공간에서 변수의 분산을 최대화하기 위해 직교 변환을 적용한다. MDS는 변환된 공간의 항목 간 거리가 원본 데이터의 근접성을 반영하도록 원 데이터 세트의 차원을 줄인다. DM은 데이터 세트의 기하학적 파라미터를 분석하여 데이터의 차원을 줄이려고 하는 비선형 기술로 DM은 원 데이터 세트에서 유클리드 거리로 측정된 포인트 간 위치를 분석하고, 변환된 공간에서 확산 거리가 원 유클리드 거리와 일치하게 축소 버전을 생성한다.

8. Transfer Learning

TL은 이전에 알려진 데이터 세트에 사용된 알려진 모델을 다른 응용에 적용하는 것으로, 직관적이지 않은 것처럼 보이지만 서로 다른 영역 간 지식 전달은 새로운 모델을 교육할 필요가 없으므로 학습 성능을 크게 향상시킬 수 있다. TL은 회귀, 분류와 클러스터링 문제에 적용될 수 있으며 사용되는 ML 기술의 유형에 제한이 없다. SON에서 TL은 자원 최적화와 결합 분류에 적용 연구되었다.

IV. SON의 기계학습 요구사항

SON을 공학적 개념으로 정의하기 위해서는 원하는 목적 달성을 위해 역동적인 순시 환경에서 자율적이고 지능적으로 적응하면서 독특하게 조직화된 행동을 보여주는 특정 생태계를 관찰할 필요가 있다. 물고기 떼, 벌레 떼, 양 떼, 동시에 번쩍이는 반딧불이, 인체의 복잡한 면역과 내분비계, 무리를 짓는 새들이 자연적으로 구성된 자가

행동의 예이다.

그 중에서 두루미 떼의 이동 상황을 관찰해 보면 다음과 같다[3]. 무리 속의 두루미 각 개체가 겪는 평균 공기 저항을 개별 비행 효율과 비교하여 최대 70%까지 최소화함으로써 전체 무리의 평균 비행 효율을 최대화하도록 비행 속성을 조정한다. 이는 그룹 비행 중에 최적의 델타 형태를 능동적으로 유지함으로써 가능한데, 여기서 중요한 다음 세 가지를 추론할 수 있다. 첫째, 리더가 없으면서 무리 속 새들 간 글로벌 정보 교환 혹은 조정이 없다. 즉, 명시적인 글로벌 신호와 중앙 제어가 없다는 것은 완벽한 확장성(Scalability)을 의미한다. 둘째, 바람과 기상 조건이 바뀌더라도 장기적인 비행 방향에서 벗어나거나 두루미들이 분리되지 않는다. 즉, 완벽한 안정성(Stability)을 의미한다. 셋째, 포식자 공격과 큰 장애물을 피하기 위해 집단 행동뿐만 아니라 개인의 시급한 실행 능력을 잃지 않는다. 즉, 완벽한 민첩성(Agility)을 의미한다.

따라서 두루미 무리는 적응적이고 자율적이며 긴급한 조직적 행동을 한다는 것을 SON에 적용한다는 것은 두루미 무리의 세 가지 특징인 확장성, 안정성, 민첩성의 Trilemma 문제를 해결해야 한다. 그러나 현재는 수학적 이론적 접근으로 불가능하므로 기계학습 기술 적용이 한 방안이 될 수 있다.

그동안 SON의 특정 유스 케이스를 해결하기 위한 지능형 솔루션 연구가 진행되었다. 적용 알고리즘으로는 기본 제어 루프와 임계값 비교에서부터 보다 복잡한 ML과 데이터 마이닝 기술이 포함된다. 각각의 SON 기능에는 고유한 요구사항이 있으므로 특정 기능에 대해 특정 ML 알고리즘이 더 잘 작동하는 경향이 있는데, 이를 구체적으로 분석하기 위한 메트릭으로 확장성, 훈련 시간, 응답시간, 훈련 데이터 양, 복잡도, 정확도, 수렴

시간과 수렴 신뢰도의 개념과 특징을 검토해 보면 다음과 같다[11].

확장성은 시스템에 더 많은 데이터를 공급하거나, 입력 데이터에 더 많은 특징을 추가하거나, 신경망에 더 많은 계층을 추가하는 등 규모 증가에 대해 무한대로 복잡성을 증가시키지 않으면서 처리한다는 것을 의미한다. 미래 네트워크는 훨씬 더 밀도가 높고 더 많은 데이터를 생성할 것으로 예상되므로 확장성은 망에서 쉽고 빠르게 알고리즘을 배포할 수 있는 매우 바람직한 특징이다. 확장성은 특정 ML 알고리즘이 분산 솔루션에 대량으로 설치될 수 있는지 또는 중앙집중식 솔루션으로 선호되는지 결정에 도움이 될 수 있다. 확장성이 필요한 SON 기능 예로는 망에서 사용자 이동성 패턴을 예측하려는 알고리즘과 결합의 감지·관리를 위해 전체 망을 모니터링하는 자가 치유 등이다.

훈련 시간은 각 ML 알고리즘이 완전히 학습하고 예측을 수행하는 데 걸리는 시간으로, 수행되는 훈련에 따라 특정 ML 알고리즘이 특정 SON 기능에 더 적합할 수 있다. 이동성 관리, HO 최적화, SON 기능 조정 또는 자가 치유와 같이 시간에 크게 의존하는 기능은 온라인 훈련이 가능한 ML 알고리즘이지만, 오프라인 훈련으로 요구 시간 내에 적용할 수 있는 ML 알고리즘도 적용 가능하다.

SON 민첩성과 관련하여 ML 알고리즘의 응답 시간으로 훈련된 알고리즘으로 원하는 SON 함수를 위한 예측에 걸리는 시간이다. 높은 응답 시간을 갖는 ML 알고리즘은 예측이 시간 내에 불가능하므로 시간 민감한 SON 기능에 적용은 바람직하지 않다. 망의 대부분 구성 파라미터가 오프라인 방식으로 결정되는 자가 구성과 같은 SON 기능에는 빠른 응답 시간을 요구하지 않으므로 낮은 응답 시간을 가진 ML 알고리즘이 적합하다. 그러

나 이동성 관리, HO 최적화, 호 수락 제어, SON 함수 조정 및 자가 치유와 같은 기능에는 더 빠른 응답 시간이 필요할 수 있으며, 더 빠른 ML 알고리즘을 적용해야 한다.

ML 알고리즘 파라미터와 관련하여 해당 알고리즘에 필요한 훈련 데이터의 양과 유형이 있다. 특정 유형의 ML 알고리즘은 레이블이 있거나 레이블이 없는 데이터에서만 작동하며, 이는 특정 유형의 SON 함수에 가장 적합할 수 있다. 많은 훈련 데이터가 필요한 ML 알고리즘은 일반적으로 정확도가 높지만 훈련에 더 많은 시간이 걸린다. 성능을 높이기 위해 많은 양의 데이터에 의존하는 알고리즘은 데이터를 수용하고 데이터를 사용하여 모델 훈련을 위해 더 많은 메모리가 필요한데, 메모리 기능 제한으로 이동성 예측, HO 최적화와 같은 사용자 단말에 설치되는 기능과 호환되지 않을 수 있다. 자가 치유의 경우 운영자는 망 모니터링과정에서 레이블이 없는 많은 데이터를 수집하는 경향이 있다.

복잡도는 원하는 솔루션 달성을 위해 수행하는 수학적 연산의 양으로 정의되며, 특정 ML 알고리즘이 사용자 또는 운영자가 설치하기에 더 적합한지 여부를 결정할 수 있다. 또한 더 복잡한 시스템은 결과를 생성하는 데 시간이 더 걸리지만, 이 결과는 다른 간단한 접근 방식보다 더 나은 경향이 있다. SON 기능으로 일반적으로 간단한 솔루션이 선호되지만 때로는 간단한 솔루션으로 충분한 결과를 얻을 수 없다. 훨씬 더 조밀하고 수천 개의 BS 내 파라미터가 예상되므로 미래 네트워크에서 자가 구성은 간단한 솔루션이 충분하지 않으며 더 복잡한 솔루션이 필요할 것이다. 반면 간단한 솔루션은 자가 치유 기능에 적합하므로 향후 망을 사전에 예방하고 오류를 감지하고 완화할 수 있다.

특정 SON 기능의 우수한 동작을 보장하려면

정확도가 높은 ML 알고리즘이 필요하다. 결합 검출에서는 망 오류를 올바르게 감지한다면 다른 SON 기능의 응답이 더 빨라지고 망 오류의 영향을 완화할 수 있다. 그러나 다른 기능은 매우 높은 정확도를 요구하지 않을 수 있으며 더 관대할 수 있다. 예를 들면 셀 커버리지 영역 추정에서 정확한 커버리지 영역이 결정될 필요가 없고, 추정치가 충분할 수 있고, 부하 밸런싱 측면에서 전체 망의 완벽한 부하 밸런싱이 필요하지 않거나 가능하지 않을 수도 있다.

ML 알고리즘 수렴 시간은 해당 특정 문제에 대해 찾은 솔루션이 당시 최적의 솔루션이라고 동의하는 데 걸리는 소요 시간이다. 컨트롤러 또는 RL과 같은 알고리즘은 솔루션이 수렴되고 다음 슬롯에서 갑자기 변경되지 않도록 추가 시간이 필요하다. 수렴 시간은 시스템 응답 시간 외에 추가 시간이 추가되므로 이 추가 파라미터가 있는 솔루션은 이동성 또는 HO 최적화와 같은 시간에 민감한 기능에서 제대로 수행되지 않을 수 있다. 그러나 솔루션이 수렴되었고 그 당시 가능한 최상의 솔루션임을 보장하는 ML 알고리즘은 시스템에 거의 최적의 솔루션을 제공할 수 있다. 이런 ML 알고리즘을 활용하는 SON 기능은 시간에 민감하지 않은 BS 캐싱과 자원 최적화의 초기 파라미터를 신중하게 조정하는 자가 구성 기능 등이다.

학습 알고리즘의 또 다른 중요한 파라미터는 설정되는 초기 조건들과 그것들의 수렴 신뢰성인데, 이 수렴 신뢰성 메트릭은 알고리즘이 로컬 최소점에 갇히게 될 민감성과 초기 조건이 성능에 미치는 영향을 나타낸다. 정확도와 관련되어 있지만, 로컬 최소점에 갇히는 영향을 최소화할 수 있는 ML 알고리즘이 보다 최적의 솔루션을 얻을 수 있으므로 이 메트릭은 ML 알고리즘이 로컬 최소점에 갇히거나 혹은 그렇지 않은 민감성을 나

타낸다.

Klaine [11]의 연구에서는 SON에 적용된 ML 알고리즘들에 대해 SON 요구사항 관점에서 성능 메트릭에 대해 현 알고리즘 관점에서 정리하였으며, 향후 개선된 알고리즘의 경우에는 다를 수 있다. 확장성 측면에서는 ML 알고리즘으로는 DT, CF, K-Means, SOM, Feedback, MC이 가장 뛰어나고, 훈련 시간 측면에서는 Bayesian, k-NN, DT가 가장 좋고, 응답 시간 측면에서는 Bayesian, k-NN, NN, SVM, DT, CF, K-Means, SOM, MC, HMM이 가장 좋다. 훈련 데이터 측면에서는 Bayesian, k-NN, DT, K-Means, QL이 가장 좋다. 복잡도 측면에서는 Bayesian, k-NN, DT, K-Means, SOM, Feedback, QL, MC가 가장 우수하며, 정확도 측면에서는 NN, SVM, CF, SOM, GA가 가장 좋다. 수렴 시간 측면에서는 현 ML 알고리즘에 대한 연구가 가장 미흡하므로 추후 평가가 필요하며, 수렴 신뢰도 측면에서는 CF, GA가 가장 좋다.

그런데, SON에 적용된 ML 알고리즘을 보면 SON 목적 실현을 위해 두 개 이상의 ML 알고리즘을 결합하여 각 ML 알고리즘의 장점을 활용하고 있으므로, Klaine [11]의 연구는 ML 알고리즘 결합을 위한 선택 시 참고할 만하다.

V. 결론

본 고에서는 이동통신망의 계획, 구성과 최적화 과정에서 전문가 개입을 최소화할 수 있는 SON 실현 기술로 ML 알고리즘을 분석 정리하였다. 우선 ML 알고리즘의 기본 개념을 간단하게 정리하였고, 이 알고리즘들이 각각 SON의 어느 영역에 적용되었는지를 간단하게 정리하였다. 그리고 SON 관점에서 ML에 대한 요구사항과 성능 메트릭을 정리하였고, 그리고 지금까지 SON에 적용

되었던 ML 알고리즘이 SON 성능 메트릭 관점에서 성능이 어느 정도인지 정리하였다.

SVM	Support Vector Machine
TCA	Trigger-Condition-Action
TL	Transfer Learning

약어 정리

BS	Base Station
CF	Collaborative Filtering
DB	Data Base
DT	Decision Trees
FLC	Fuzzy Logic Controller
GA	Genetic Algorithm
HetNets	Heterogeneous Networks
HMM	Hidden Markov Model
HO	Hand-Over
k-NN	k-Nearest Neighbor
KPI	Key Performance Indicator
MC	Markov Chain
MCA	Minor Component Analysis
MDS	Multi-Dimensional Scaling
ML	Machine Learning
MLB	Mobility Load Balancing
MRO	Mobility Robustness Optimization
NCL	Neighbor Cell List
NN	Neural Network
PCA	Principal Component Analysis
QL	Q-Learning
RL	Reinforcement Learning
SO	Self-Organizing
SOM	Self-Organizing Map
SON	Self-Organizing Network

참고문헌

- [1] R. Barco, V. Wilee, and L. Diez, "System for Automated Diagnosis in Cellular Networks based on Performance Indicators," *Eur. Trans. Telecommun.*, vol. 16, no. 5, 2005, pp. 399-409.
- [2] J. Laiho et al., "Advanced Analysis Methods for 3G Cellular Networks," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, 2005, pp. 930-942.
- [3] O. G. Aliu et al., "A Survey of Self-Organisation in Future Cellular Networks," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, 2013, pp. 336-361.
- [4] S. Bi et al., "Wireless Communications in the Era of Big Sata," *IEEE Commun. Mag.*, vol. 53, no. 10, Oct. 2015, pp. 190-199.
- [5] M. Peng et al., "Self-Configuration and Self-Optimization in LTE-Advanced Heterogeneous Networks," *IEEE Commun. Mag.*, vol. 51, no. 5, May. 2013, pp. 36-45.
- [6] R. Barco et al., "A Unified Framework for Self-Healing in Wireless Networks," *IEEE Commun. Mag.*, vol. 50, no. 12, Dec. 2012, pp. 134-142.
- [7] A. Tall et al., "Distributed Coordination of Self-Organizing Mechanisms in Communication Networks," *IEEE Trans. Contr. Netw. Syst.*, vol. 1, no. 4, Dec. 2014, pp. 328-337.
- [8] I. Karla, "Resolving SON Interactions Via Self-Learning Prediction in Cellular Wireless Networks," in *Proc. Int. Conf. Wireless Commun. Netw. Mobile Comput.*, Shanghai, China, Sept. 2012, pp. 1-6.
- [9] H. Y. Lateef, A. Lmran, and A. Abu-dayya, "A Framework for Classification of Self-Organising Network Conflicts and Coordination Algorithms," in *Proc. IEEE Annu. Int. Symp. PIMRC*, London, UK, Sept. 2013, pp. 2898-2903.
- [10] P. Wainio and K. Seppanen, "Self-Optimizing Last-Mile Backhaul Network for 5G Small Cells," in *Proc. IEEE Int. Conf. Commun. Workshops*, Kuala Lumpur, Malaysia, May. 2016.
- [11] P. V. Klaine et al., "A survey of machine learning techniques applied to self-organizing cellular networks," *IEEE Commun. Surveys Tutorial*, vol. 19, no. 4, July. 2017, pp. 2392-2431.