
Global Report

건물 설계에서의 빅데이터: 고찰

1. 개요
2. 정의
3. 적용
4. 논의
5. 결론

1. 서론

많은 저자들이 빅데이터를 지난 반세기 동안 가장 중요하고 혁명적인 과학 및 산업 현상 중 하나라고 주장한다 (Walker 2014). 소비자 행동 예측에서부터 은행 사기 탐지, 항공편 스케줄링 및 맞춤형 의료 치료에 이르기까지 다양한 어플리케이션들이 존재한다. 실제로 인간 활동의 거의 모든 영역인 비즈니스(Chen 외. 2012, Minelli 외 2012), 경영(Ju 외. 2013), 금융 (Fang & Zhang 2016), 인적 자원 관리(Biela & Cochran 2014), 건강 관리(Bellazzi 2014, Groves 외 2013), 공공 행정 (Maciejewski 2016, Martin 2014), 관광 (Xiang 외 2015, Fuchs 2014), 법 집행(Glasgow 2015), 그리고 과학의 모든 영역(Bryant 2011, Szalay 2011, NIST 2015b)에서 빅데이터의 활용이 문서화되고 있다.

건축, 엔지니어링, 건설 및 운영(이하 AECO; architecture, engineering, construction and operation) 산업도 예외는 아니지만 현재의 개발 수준은 상대적으로 훨씬 낮다. AECO 분야에서 같은 양의 데이터도 수집되지 않으며 같은 강도(intensity)로 사용되지도 않는다. 학계에서 실험 목적으로 주로 실행될 수 있는 구체화된 어플리케이션들조차 없다. 실습을 위한 양도 가능성도 여전히 불투명하다. 그러나 상당수의 학자와 실무자들은 빅데이터의 사용이 건축 산업에 상당한 영향을 미칠 것으로 생각한다 (Bilal 외 2016, Deutsch 2015a, Deutsch 2015b). 빅데이터가 전례없는 수준의 디테일과 특수성을 갖는 디자인을 설계자에게 알려줄 뿐만 아니라 전통적으로 주관적이고, 매우 복잡하거나 예측할 수 없는 현상을 모델링할 수 있도록 해 준다는 점은 명백하다.

이웃의 "캐릭터"를 평가하거나 건물 외관의 "매력도"를 측정하는 것에서부터 행동 패턴을 규명하거나 사무소 공간의 점유 일정을 예측하는 것에 이르기 까지 기계 학습 알고리즘과 결합된 빅데이터는 건물 설계 과정에서 이해하고 시뮬레이션하고 측정할 수 있는 것들을 재정의하는 것처럼 보인다.

이 논문에서 우리는 건물 설계에서 빅데이터 활용을 위한 기술, 어플리케이션 및 기회에 대해 포괄적이고 체계적으로 고찰한다. 첫째, 우리는 이 기술 현상을 형성하는 주요 개념과 기술에 대한 전반적인 견해를 제시하고자 한다. 다음으로 빅데이터 및 기계 학습 응용 프로그램을 약 100건의 건물 설계 적용 사례들을 조사하여 파격적인 12개 부문에 대해 규명한다. 마지막으로, 우리는 더 많은 연구가 필요한 3가지 초기 분야에서의 기술 영향력에 대해 논의하고자 한다. 첫째, 어떻게 빅데이터가 점유자, 컨텍스트 및 기타 관련 에이전트의 상세한 특성을 제공하여 설계 문제를 보다 철저하고 정확하게 정의하는지, 둘째, 오직 기계 학습을 통해 개발된 통계 모델을 기반으로 오늘날 과학적 설명 모델 부족으로 예측할 수 없는 것들을 포함하여 설계물의 성능에 대한 복잡한 시뮬레이션을 수행할 수 있는지, 셋째, 보다 복잡하고 다양한 변수의 측정기준을 포함하도록 건물 성능의 개념을 재검토하면서 전(全) 생애주기 동안 프로젝트와 건축 설계물을 지속적으로 평가할 수 있는 기회에 관해 논의하고자 한다.

2. 정의

2.1 건물 설계

이 논문에서 건물설계(building design)는 건물의 시공과 관리에 필요한 모든 정보의 개념화, 공식화, 정교화를 의미한다. 건물 "설계"는 건축가에 의해 수행되는 예술적 행위뿐만 아니라 기술적 행위이지만 다양한 분야의 전문가들의 참여가 필요한 통합된 행위이다.

2.2 데이터 기반 설계

건축 설계에서 데이터 사용은 새로운 것이 아니다 (Oechslein 1993; Schmitt 1999). 클라이언트 브리핑 내용, 구조 계산 및 건축 법규 등은 설계를 발전시키거나 제한하는 데이터로 간주된다. 최근 AECO 산업의 추세는 의사결정 과정에서 객관적이고 검증 가능하며 정량적인 데이터를 활발하게 사용해 왔다. 다른 산업 (Karakiewicz 2010, Yin 외 2014)에서 사용된 개념을 빌리자면 이러한 추세는 데이터 기반 설계로 불린다 (Deutsch 2015a). 가장 잘 알려진 데이터 기반 설계 방법은 프로젝트

설계의 중요한 결정을 내리는 데 있어 현재의 최선 증거를 사용하는 증거 기반 설계 (evidence-based design, EBD)이다 (Watkins 2009). EBD는 사회 및 자연 과학의 전통적인(즉, 학술적인) 연구 형식으로 가공된 정보로 이해되는 데이터 활발하게 사용한다 (Brandt 외. 2010; CHD 2016).

2.3 빅데이터

빅데이터를 정확하게 정의할 수는 없으나 상업적 맥락에서, 일반적으로 시간과 자원의 허용 범위 내에서 전통적 방식으로 분석이 불가능한 데이터 집합으로 정의된다. 학문적으로는 빅데이터를 3가지 특성(크기, 다양성, 속도)을 충족시키는 광범위한 데이터 집합으로 정의한다.

2.4 기계 학습 (머신러닝, Machine learning)

기계 학습은 일반적으로 컴퓨터에 명시적인 프로그래밍 없이 어떤 대상에 대한 학습 능력을 제공하는 일종의 인공 지능으로 이해된다. 기계 학습 알고리즘은 공식적 관계 모델의 형식으로 사전 지식 없이 대량의 데이터를 검토하고 기본 패턴이나 관계 규명을 위해 통계 도구를 사용하였다. (Alpaydin 2014; Shalev-Shwartz & Benx 2014; Marsland 2015). 빅데이터 분석에는 기계 학습 알고리즘을 사용해야 한다는 아이디어가 일반화되었으나 이는 완벽하지 않다. 데이터가 구조화되고 연구자/설계자가 그들 간의 공식적 관계 모델을 정의하기에 충분한 사전 지식이 있다면, 전통적인 통계 기법이 적합하다. 하지만 대부분의 빅데이터가 이러한 조건들을 충족시키지 못하므로 구조화되지 않은 불명확한 데이터를 처리하는 방법에 의존할 수 밖에 없다. 데이터와의 상호 작용 유형 (Ayodele 2010)에 따라 기계 학습 알고리즘은 다음과 같이 분류할 수 있다.

- 지도 학습(supervised learning): 목표가 입력 → 출력의 쌍으로 매핑된 레이블이 지정된 데이터 세트를 분석하고 이전에 없던 값들을 매핑할 수 있는 일반 함수를 추론하는 경우
- 비지도 학습(unsupervised learning): 목표가 레이블이 지정되지 않은 데이터 세트를 분석하고 데이터 구조(일반적으로 그룹이나 클러스터)를 설명하기 위한 함수를 추론하는 경우
- 강화 학습(reinforcement learning): 목표가 외부 보상 피드백의 최대화에 기초한 특정 상황에서 이상적인 값이나 행태를 자동으로 결정하는 경우

건축 설계 시 사용되는 가장 일반적이고 영향력있는 알고리즘 중 일부는 다음과 같다 (Wu et al. 2008; Lee & Siau 2001).:

- 의사 결정 나무(Decision trees)는 몇몇 입력 값들(가지)를 기반으로 출력값(잎)을 예측하는 모델을 구축하는 지도 학습 알고리즘이다.
- 출력값이 클래스인 경우에는 분류 나무(classification tree)라고 하고 출력값이 실수인 경우에는 회귀 나무(regression tree)라고 한다.

따라서, 의사 결정 나무 알고리즘을 CART(classification and regression trees)라고 부른다. 대개의 의사 결정 나무 알고리즘들은 하향식 방식에 따라 구성되어 각 단계(가지)에서 데이터 세트를 가장 잘 나누는 변수를 선택한다. 어떤 문제에서는 여러개의 상이한 의사 결정 나무들이 사용되기도 하는데 이로한 기술은 '무작위 숲'(랜덤 포레스트, random forests)으로 불린다 (Rokach & Maimon 2014).

인공 신경망 (ANN, artificial neural networks)은 구조화 및 구조화되지 않은 학습에서 많은 어플리케이션들과 함께 점차 인기를 얻고 있는 알고리즘이다. 생물학적 신경망에서 영감을 얻은 이들은 비선형 방식으로 데이터를 처리하고 분석하는 반복적인 "인공 뉴런"으로, 반복적인 결과(즉, 학습)에 따라 가중치를 지속적으로 업데이트한다. 인공 신경망은 입력과 출력 사이의 복잡한 관계들을 모델링하고 데이터 내에서 예기치 않은 패턴이나 구조를 찾아내는 등 다양하게 활용되고 있다. 입력 → 출력 뉴런만 있는 단순한 단일 수준의 인공신경망과 달리 여러 층의 인공 뉴런을 갖고 있는 인공신경망은 '딥러닝 인공신경망'으로 불린다.

서포트 벡터 머신(SVM, Support Vector Machine)은 분류 및 회귀에 사용되는 지도 학습 알고리즘

으로 선형 분류기, 즉 유사한 특징을 갖는 선형 조합의 값에 따라 그룹을 분리하는 알고리즘이다 (Ayodele 2010). 레이블이 지정된 데이터 세트로 부터 SVM은 새로운 사례들을 분류하는 데 사용되는 최적의 n 차원 초평면(hyperplane)을 얻는다. 어떤 사례가 특정 클래스와 유사한 지에 대한 여부를 감지하는 데 주로 사용된다.

군집 분석은 특정 기준에 따라 일련의 값을 등급 또는 군집으로 분류하는 비지도 학습 수단이다. 다양한 기준이 서로 다른 군집화 알고리즘을 발생시킨다. 가장 일반적인 알고리즘 중 하나는 k -평균으로, 공간을 각 군집이 일종의 다차원 보로노이 셀(Voronoi cells)에 가장 근접한 평균을 가진 군집에 속하는 k 개의 군집으로 나눈다. 이 알고리즘의 주요 약점들 중 하나는 특이치에 매우 민감하다는 것이다.

유전 알고리즘(GA, genetic algorithm)은 기술적으로 자연 선택 과정을 기반으로하는 메타휴리스틱(metaheuristic) 검색 알고리즘이다. 모집단과 최적화 문제(목적 함수)가 주어지면 알고리즘은 평가된 사례들로부터 무작위 모집단을 샘플링하고 기본 매개 변수("유전자")에 기초하여 확률적으로 그들을 재결합한다. 새로운 세대는 새로운 무작위 샘플을 추출하는 데 사용된다. 여러번의 반복 ("세대")이나 목적 함수의 임계값에 도달할 때까지 과정이 반복된다. 유전 알고리즘은 구현이 매우 간단하지만 전체적인 최적값을 얻을 수 없고 부분적 최적화에 수렴하는 경향이 있어서 매우 제한적이다. 이 알고리즘은 구조가 불분명한 아주 넓은 검색 공간 내에서 근사 해법을 찾는 데 사용된다. 건물 설계에서 기하학적 문제의 최적화 알고리즘으로 매우 유명하다.

2.5 가시화

데이터 분석 과정은 정량적인 성격을 띄므로 이 방법에 익숙하지 않은 사용자들에게, 특히 설계 맥락에서 문화적 장벽이 생기는 경우가 많다. 따라서 빅데이터로 작업하는 과정에서 매우 중요한 단계는 해석을 용이하게 하기 위한 데이터의 통신과 가시화이다. 건물 설계 분야, 특히 공간 정보 측면에서 데이터의 가시화는 중요한 역할을 한다. 데이터 가시화, 데이터 분석 및 인간 - 컴퓨터 상호 작용을 결합하여 추론, 이해 및 의사 결정을 용이하게하는 방법을 시각적 분석(visual analytics)이라고 한다 (Keim 외. 2010). 수많은 디지털 시각적 분석 도구가 있는데 일부는 오픈 소스이지만 높은 수준의 전문 지식을 필요로 한다. 프로그래밍 언어 R은 특히 전용 그래픽 패키지 (예 : gplot)와 함께 데이터 분석에 가장 일반적으로 사용되는 환경 중 하나가 되었다. 다른 소프트웨어 솔루션(예 : Tableau, Spotfire, QlikView)은 상업용이고 사용이 매우 간편하지만 일반적으로 분석의 유연성이 제한적이다 (Zhang 외. 2012).

3. 적용

3.1 사전 설계

사전 설계 단계는 설계 프로세스의 첫 단계로 이 단계에서 설계 문제가 필요한 것(범위, 목표, 요구)과 어떤 조건(사용 가능한 리소스, 제한, 제약)의 견지에서 정의된다 (Hershberger 2015). 빅데이터가 유용한 두 가지 중요한 사전 설계 작업은 기획(programming)과 대지 분석이다.

3.1.1 기획

기획은 설계 프로젝트의 요구 사항들을 식별하고 연구하는 것으로 이 작업의 범위는 프로젝트의 규모와 복잡성에 따라 다르지만 가장 전통적인 형식으로 필요한 공간의 정의, 기능, 예산, 기간 및 시공 조건을 정의하는 것이다 (Kumlin 1995).

이 단계에서 주요 데이터의 근원은 전형적으로 반드시 객관적이고 정확하고 충분한 것이 아닌 클라이언트의 요구들에 대한 자각이므로 수 많은 기획 지침은 설계자들이 추가적인 정보를 수집할 것을 적극 권장했다. 이 정보는 매우 유용할 테지만 얻기가 어렵고 비용이 많이 든다. 하지만 이 정보를 제공할 수 있는 빅데이터의 몇가지 출처들이 있다. 예를 들면 소셜 네트워크와 지역화된 장치들로부터

터 얻어진 데이터가 개인과 집단의 수준에서 행태와 공간 이용 선호도를 찾아내기 위해 사용될 수 있다.

데이비스(Davis, 2016)는 온라인 앱으로 건물 거주자로부터 수집된 데이터를 새로운 사무실 회의공간 설계 요구조건들을 정의하기 분석한 사례들을 보여주었고 Sokmenoglu 등(2010)은 사전 공개 설문 조사에서 얻은 데이터로 이스탄불의 특수 학생 커뮤니티의 사회-공간 활동의 패턴과 추세를 조사했다. Chen와 Chiu (2006)는 가로 설계와 도시 시설과 관련된 보행자 배회 패턴을 수집하는 시스템을 제안했다. Yan (2006)은 서로 다른 건축 공간에 있는 사람들의 행동 패턴을 연구하기 위해 자동화된 비디오 추적 기반 시스템을 개발했다. Ugarte와 Leef (2016)는 주어진 영역에서 근접 관계의 개념을 정성적으로 평가하기 위해 인스타그램(Instagram) 이미지를 기반으로 클러스터링 알고리즘을 설명했다. Friedrich(2013)는 디자인을 알려주는 트렌드를 찾는 미래의 건물 사용자의 의견을 대규모로 집계하고 분석하기 위해 대규모로 수집된 소셜 네트워크와 데이터를 사용하는 방법을 보여주었다. 온라인 네트워크와 빅데이터 시스템이 방대한 인구 집단에 대한 참여적 설계를 가능하게 하여 군중 기반의 설계방법의 촉매가 될 것이라는 것이 예측된다 (Williams 과 Sotta 2015).

다른 산업 분야에서 빅데이터의 주요 응용 분야 중 하나는 관련자들(예: 고객, 환자)을 연구하고 특성화하는 것이다. 신용카드나 공개 온라인 네트워크에서 수집된 데이터는 시장을 세분화하고 개별화된 마케팅 캠페인을 생성하는 데 널리 사용된다 (Grigsby 2016; Adomavicius & Tuzhilin 2001; Linoff & Berry 2001; Artun & Levin 2015). 클라이언트와 거주자들의 선호도와 요구 사항을 이해하기 위해 유사한 방법이 건축설계에서 사용될 수 있다.

3.1.2 대지분석

사전 설계 단계에서 또 다른 중요한 작업은 건물 대지를 분석하는 것이다. 설계자들이 갖고 있는 정보가 많을수록 그들이 대지의 제약과 기회를 더 잘 판단할 수 있다. 설계자가 수집한 일반적인 정보는 물리적 데이터(예: 지형 조사, 지질학 연구), 환경 데이터 (예: 기후 차트), 법률 판례 (예: 구역화, 규제), 인프라 데이터 (예: 공공 서비스, 도로 접근성, 교통 고려 사항)를 비롯해서 심지어 대지에서의 우세한 전망이나 주변 환경의 특수한 문화적 고려 사항등과 같은 기술적이지 않은 대지 데이터도 포함된다 (LaGro 2013, Zimmerman 2000). 이 정보는 일반적으로 서로 다른 출처, 즉 조직화되지 않았거나 불완전한 상태로 분산되어 있어서 검색과 사용이 어렵다. 설계자들은 빅데이터와 지리 정보 시스템(GIS)을 결합하여 대지에 대한 방대한 지형 공간 정보를 수집, 구성, 해석하여 기존 소스를 훨씬 뛰어넘는 GIS 분석을 확장시켜서 편리한 해법을 찾는다. 예를 들어, CitysDK는 대규모 공공 정부 소스를 기반으로 작성된 네덜란드의 건물 온라인 지도이다(WAAG & Span 2016). 포괄적이지 않지만 전문화된 Historic Quest DC는 개별 건물의 역사 및 기술적 데이터를 포함한 12만 7천개 역사적 건물들의 지도이다 (HPO 2016). 빅데이터와 GIS의 결합은 범죄(Igarapé Institute 2016)로부터 가구 소득 (ESRI 2016)에 이르기까지, 재산(Petroutsos 2014)으로부터 가로별로 가장 많이 사용하는 공통 언어(Hubley 2016)에 이르기까지 거의 모든 지리적 속성을 분석할 수 있게 해준다. 우수 사례는 MIT GeoWeb으로 다양한 공공 및 학술 자료 (MIT GIS Services 2016)에서 2000개 이상 계층의 GIS 데이터 저장소를 검색, 확인 및 다운로드 할 수 있는 온라인 도구이다.

3.2 개요 설계 (Schematic design)

개념적 또는 개요적 설계 단계에서 규모, 형태, 공간의 크기와 구성, 건물의 일반적 이미지 및 공사비의 예비 추정치를 포함하여 설계물의 일반적인 특성이 정해진다(AIA 2014).

3.2.1 생성적 설계 탐색

설계 프로세스의 초기 단계는 설계자가 설계 문제를 더 잘 이해하고 예기치 않았던 해법을 찾을 수 있도록 돕는 반복적 과정 중에 신속하게 평가되고 재생성되는 여러 설계안들의 신속한 생성 작업을 수반한다. 수많은 연구자들이 중요하지만 시간 소모적인 작업을 지원할 생성적인 설계 도구를 개발해야 하는 도전에 직면하고 있다. 검색 알고리즘을 사용하여 조합적인 해법을 시도했지만 그 결과물

을 쉽게 얻어내지 못했다. 가장 단순한 설계 문제에 대해서조차도 모든 가능한 설계 옵션들의 탐색 공간이 너무나 방대해서 전산 분석을 할 수 없기 때문이다(Evins 2013). 다른 방법은 빅데이터 분석 도구를 사용하여 상당한 수의 조합적 설계물(모든 조합은 아님)을 분석하고, 통계적으로 학습한 특정 특성을 충족시키는 새로운 해법을 생성하거나 검색하는 것이다. 예를 들어, Merrell 등(2010)은 주어진 고급 요구 조건들을 고려하여 실제 데이터에 대해 학습된 베이지안(Bayesian) 네트워크를 사용하여 공간들의 목록과 이들 간의 관계를 종합한 다음 확률적 최적화를 통해 얻어진 평면을 생성하는 시스템을 제안하였다. Michalek 등(2002)은 진화 알고리즘을 사용하여 수많은 설계물을 검색하여 사각형 1층 아파트에 대한 자동 배치를 하는 방법을 제안했다. Mehanna (2013)는 기계 학습과 진화 최적화를 기반으로 한 개념 구조 설계 시스템을 제안했습니다. Erhan 등(2014)은 설계할 공간을 관리 가능한 범위로 반복적으로 축소하는 유사성 기반 검색 알고리즘을 기반으로 개념적 설계 탐색 방법을 제안했다.

3.2.2 설계 검증

설계자는 다중적 설계 옵션 생성과 더불어 설계물들이 기획 단계에 명시된 요구조건들을 충족하는지 확인해야 한다. 많은 상충되는 요구조건들을 가진 복잡한 프로젝트의 경우에 이는 자동화된 규칙 검사 시스템들이 보조해하는 어려운 작업이 된다. 이러한 시스템들은 공간 검증, 교통 문제(길 찾기), 인체 공학 연구, 시공 조건 등과 같은 기타 여러 요구조건들을 점검하는 데 점점 더 많이 사용되고 있다 (Solihin & Eastman 2015, Nawari 2012, Pauwels 외. 2011).

3.2.3 전례 분석 (Precedents Analysis)

전례에 대한 주요 출처는 전문적 경험을 제외하고 전문 저널과 웹 사이트이다(Clark & Pause 2012). 예를 들어, 가장 큰 건축 온라인 데이터베이스인 ArchDaily에는 약 20,000 개 건물의 이미지가 30만 개 이상 포함되어 있으며 매일 1,000만 명 이상의 방문자들이 문한다. (Morales 2014, ArchDaily 2016). 구조 기술 프로젝트의 온라인 데이터베이스인 Structurae는 68,000개의 구조에 대한 데이터를 제공한다(Structurae 2016). 그러나 이러한 웹 데이터베이스들은 풍부한 콘텐츠를 갖고 있지만 조직적으로 열악하다. 관련된 전례를 찾는 것은 느리고 노동 집약적인 과정이다. 몇몇 연구자들은 건물을 분석하고, 분류하고, 의미 정보를 추출하는 과정을 돕는 빅데이터 기반 시스템을 제안했다. 건물 설계 정보의 대부분은 이미지 형식 (예 : 평면도, 그림, 사진)으로 저장되기 때문에 이러한 시스템들은 일반적으로 공간 구성, 공간 구조 또는 건축 양식과 같은 의미있는 기능을 탐지하기 위한 방법들로 구성된다. 예를 들어, Lin (2011, & Chiu 2010)은 k-means 클러스터링 방법을 사용하여 평면도의 공간 위상 패턴을 자동으로 대량 분류할 수 있는 시스템을 개발했다. Strobbe 등 (2016)은 특정 건축 양식에 속한 건축 설계물의 자동 탐색에 관한 문제를 해결했다. Kuo (2003)는 중국의 전통 건축 서적에서 텍스트 데이터를 채취하여 중국 정원 설계를 분석하는 계획을 제안했고 Reffat (2008)는 사우디 아라비아의 현대 건축 패턴을 파악하기 위해 공공 데이터를 얻는 방법에 관해 논의했다.

3.3 설계물 개발과 시공 도서 작성

3.3.1 건물 데이터 관리

AECO 산업의 주요 데이터 관리 기술은 빌딩 정보 모델링(BIM), 즉 건물의 전생애주기에서 걸쳐서 나타나는 모든 물리적 및 기능적 특성들의 통합된 객체 지향적 데이터베이스이다 (Eastman 2011). 그러나 실무에서 BIM 기술은 잠재적인 용도와 관련하여 거의 활용되지 못하고 있다. 모든 참가자가 프로젝트의 전체 생애주기 동안 모든 건물 데이터의 통합 모델을 사용할 때 BIM의 가장 큰 이점이 달성된다는 사실이 널리 받아 들여지고 있다. 이러한 우수한 활용은 영국, 미국, 싱가포르, 핀란드, 노르웨이 및 덴마크를 비롯한 일부 국가에서만 문서화되었다. 다른 대부분의 국가에서 BIM은 아직 초기 단계에 있으며 시각화, 시공 문서 제작, 기본 시스템 조정 및 물량 산출에 초점이 맞춰져 있다 (McGraw-Hill Construction 2009, 2010, 2012a, 2012b, 2014a, 2014b, Loyola 2016, 2018, NBS 2018). 몇몇 연구자들이 더 많은 데이터로 BIM 모델을 증강시키기 위해 빅데이터와 기계 학습을 기

반으로하는 틀과 방법을 제안했다. Krijnen과 Tamke (2015)는 IFC의 기하학적 데이터를 기반으 기계 학습 방법을 사용하여 BIM 모델에서 암시적 지식을 평가하는 방법을 보여주었다. Lan과 Chiu (2005)는 데이터 마이닝 기술을 적용하여 협업 설계 프로젝트에서 작업 팀들의 소통 정보에서 설계물의 의미론적 패턴을 발견했다. Chen 등(2016), Amarnath 등(2011), Chuang 등 (2011), Jiao 등(2013)은 온라인에서 호스팅되는 BIM 모델을 상호 연결하는 다양한 방법과 가능성을 모색하여 센서의 연속적인 실시간 데이터를 사용하여 대규모 협업 정보 시스템을 구성했다.

3.3.2 에너지 연구

지금까지 에너지 연구는 빅데이터가 가장 성공적으로 활용된 건물 설계 분야에서 나타난다. Magoulès & Zhao (2016)는 건물 에너지 분석을 위한 기계 학습 방법에 대한 포괄적인 검토물을 제공했다. Foucquier 등(2013)은 물리적 모델과 기계 학습 모델을 비교하면서 건물 모델링 및 에너지 성능 예측과 관련된 최첨단 기술을 검토했다. Zhao와 Magoulès (2012)는 건물 에너지 소비 예측을 위해 최근에 개발된 통계 모델과 인공 지능 방법을 검토했다. Yu 등 (2013)은 건물 에너지 데이터에서 지식을 추출하기 위해 일반적으로 사용되는 데이터 마이닝 방법을 검토했다. Borgstein 등(2016)은 건물 에너지 성능을 분석, 분류, 벤치마킹, 등급화, 평가하기 위한 유용한 방법들을 검토했다. Kalogirou (2000, 2001)는 재생 에너지 시스템의 설계 및 평가에 인공 신경망의 적용을 검토했다. 특히 설계 최적화에 빅데이터 활용과 관련하여 Evins (2013)는 지속 가능한 건물 설계에 적용되는 전산 최적화 방법들을 검토했다. Kalani 등 (2016)은 데이터 마이닝을 통한 건물 에너지 소비 최적화에 대한 다양한 연구를 검토했으며 Machairas 등(2014)은 건물 설계 최적화를 위한 몇몇 알고리즘을 검토했다.

일반적으로, 에너지 분석에서 빅데이터와 기계 학습은 사용할 수 있는 물리적 모델이 정립되지 않는 현상의 분석, 시뮬레이션 및 예측에 점점 더 많이 사용되고 있다. 빅데이터와 기계 학습은 통계 모델을 기반으로 운영되므로 열역학 방정식, 조명 거동 가설(lighting behavior assumptions), 음향 속성 데이터 또는 기하학적 매개 변수가 필요없다. 따라서 에너지 요구를 모델링하거나, 사용자 행동을 예측하거나, 전통적인 물리적 모델을 만드는 것이 너무 복잡하거나, 부정확하거나, 비현실적이거나, 심지어 불가능한 복잡한 다중 변수 설계를 최적화하는 데 적합하다.

예를 들어, Castelli 등. (2015)은 기하학적 시멘틱 유전자 프로그래밍으로 불리는 부류의 진화 알고리즘을 사용하여 주택의 에너지 수요 예측 모델을 제안했다. Capozzoli 등 (2015)은 난방 및 가정용 온수의 1차 에너지 수요를 감지하고 가장 큰 영향을 미치는 요인을 식별하기 위해 약 9만 가구의 데이터를 분석했다. Tsanas & Xifara (2012)는 주거 지역의 난방 및 냉방 부하에 대한 물리적 변수의 영향을 연구하고 예측하기 위해 무작위 숲(랜덤 포레스트)를 기반으로 한 통계적 기계 학습 프레임워크를 개발했다. Kalogirou (2000)는 쉽게 측정할 수 있는 물리적 특성을 기반으로 패시브 태양열 건물의 에너지 소비 예측을 위해 인공신경망(ANN)을 활용했다. Yu 등 (2010)은 건물 에너지 사용 강도를 모델링하여 주거용 건물 에너지 성능 지수를 추정하기 위한 의사 결정 나무(decision tree)를 제안했다.

Li 등(2011)은 중국의 사례 연구에서 에너지 소비를 예측하기 위해 유전자 알고리즘과 적응 네트워크(adaptive network)의 하이브리드 시스템을 제안했다. 김현주 등 (2011)은 건물 설계안 개선을 위해 다른 시스템 간의 상호 관계 규명을 위한 데이터 마이닝 기술을 제안했다. Su & Yan (2014)은 유전자 알고리즘을 사용하여 레이아웃과 채광 성능을 최적화하여 기존 방식에 비해 전산 소요 시간을 효과적으로 줄였다. Feng 등 (2014)은 군집 분석과 관련 규칙을 사용하여 자연 환기 전략 및 건물 설계 권장 사항과 관련된 유효한 창문의 거동/ 작동 패턴을 식별하는 모델을 제안했다.

고급 전산 에너지 시뮬레이션은 고도의 계산 요구조건들이 필요로 해서 설계 과정 중에 병목 현상을 일으킬 수 있다. Chatzikonstantinou와 Sariyildiz (2016)는 시뮬레이션 기반 데이터로 훈련된 피드 포워드 네트워크(feed-forward networks), SVM 및 랜덤 포레스트 기반의 사무실 공간의 시각적 쾌적성을 위한 예측 모델을 구축하여 대안적 접근방법을 모색했다.

빅데이터와 기계 학습의 사용은 비교적 최신 방법이므로 각 방법의 장단점에 대해서는 아직 명확하게 밝혀지지 않았다. 여러 중요한 비교 연구들이 수행되었다. Mateo 등 (2013)은 건물 온도 예측 과정에서 다양한 분석 방법들(자가 회귀 모델, 강력 다중 선형 회귀, 다층 퍼셉트론, 엑스트림 학습기, 비선형 자기 회귀 외인성, k-평균, 퍼지 c-평균, 누적 계층 트리, DBSCAN)을 고전 기법과 비교했다. Neto와 Fiorelli (2008)는 건물 에너지 소비 예측 측면에서 인공 신경망 기반 모델을 실제 물리적 모델과 비교했다. Fonseca 등(2013)은 사무실 건물의 최종 전기 에너지 요구사항들을 구축하는데 주광(daylighting)의 영향을 예측하기 위해 인공 신경망 사용을 다변수 선형 회귀 분석과 비교했다.

3.3.3 구조 설계

구조 설계는 데이터 기반 작업으로 인식되어 빅데이터 분석법의 사용은 다른 건물 설계 분야 보다 더 자연스럽게 보인다. 유명한 적용 사례는 SHM (Structural Health Monitoring) 즉, 손상을 감지하고 특성화하기 위한 데이터 수집을 위해 실제 구조물에 센서를 활용하는 것이다. Farrar와 Worden (2012) 연구에서 기계 학습 관점에서 SHM을 다룬 내용을 찾아볼 수 있다. 에너지 연구와 유사하게 빅데이터는 실물과 모의 구조물의 대규모 데이터 집합에 대한 통계 학습을 기반으로 물리적으로 불완전한 현상을 분석하는 데 매우 유용하다. 예를 들어, 몇몇 연구자들은 복잡하고 불확정한 구조를 모델링하는 알고리즘을 개발했다. Mehanna (2013)는 경량, 가변 및 비고정 구조물을 교육하는 시스템 개발을 위해 빅데이터와 ANN을 사용하였다. Yamamoto 등(2011)은 초기 구조 구성물에 대한 최소한의 지식만으로 원하는 구조물들을 얻어낼 수 있도록 설계자를 돕는 유전 알고리즘을 제안했다. 구조 설계는 다른 건축 설계 분야보다 전문 지식 체계화가 취약한 문제들을 공유하는데 빅데이터가 문제 해결에 중요한 역할을 할 수 있다. 예를 들어 Dolšák과 Novak (2011)은 구조적 지식을 체계화해서 지능형 의사결정 지원 시스템 개발을 위한 틀을 제안했다. Freischlad 등(2006)은 철근 콘크리트 구조 부재 설계를 위한 퍼지(fuzzy) 규칙들의 데이터 추출을 위한 유전 알고리즘을 제안했다.

3.4 시공

3.4.1 시공성 분석

건물 설계에서 시공 데이터의 가장 인습적인 활용은 시공성 검토가 대표적이다. 시공성 (constructability 또는 buildability)는 1980년대에 창안된 개념으로 클라이언트의 모든 요구사항 (CIRIA 1983, CII 1986)에 따라 [건물] 설계가 시공을 용이하게하는 정도를 뜻한다. 프로젝트의 시공성 향상을 위해 설계자는 건설 프로세스의 모든 특수성을 고려해야한다. 이는 일반적으로 시공자나 경험이 풍부한 설계자에게만 있는 지식 유형이다.

몇몇 연구자들은 구조화되지 않은 원천으로 부터 시공 지식을 습득하고 구조화된 시공성 규칙을 생성하기 위해 빅데이터와 기계 학습 방법을 이용하는 시스템을 제안했다. 예를 들면, Skibniewski 등 (1997)은 31건의 건설 프로젝트 예제 데이터베이스를 이용하여 의사 결정 규칙 형식으로 자동화된 시공성 지식 습득에 대한 타당성 조사를 수행했다. Freischlad와 Schnellenbach-Held (2005)는 퍼지 모델링과 다목적 진화 알고리즘을 사용하여 구조적인 설계 지식을 수집하고 표현하는 시스템을 제안했다.

3.4.2 시공 엔지니어링

시공 엔지니어링은 시공 현장 모니터링에서부터 작업자의 부상을 예측하는 데 이르기까지 빠르게 변하는 데이터의 흐름에서 빅 데이터 및 기계 학습의 광범위한 적용을 보여주는 좋은 사례이다. 예를 들면, Lu 등(2015)은 서로 다른 범주의 프로젝트들에 대한 폐기물 발생률 지표 및 벤치 마크를 구축하기 위해 2년간 수천 개의 프로젝트에서 2백만 건 이상의 폐기물 처분 기록을 조사하였다. Asadi 등(2015)은 시공 조달 물류의 지연을 예측하기 위해 의사 결정 트리(Naive Bayes 알고리즘)를 사용했다. Yang 등(2015)은 빅데이터 및 기계 학습을 사용한 몇가지 건설 모니터링 방법을 검토했다. Li

등(1999)은 기계 학습 및 유전 알고리즘을 사용하여 시간 - 비용 거래 문제를 해결했습니다. Million 등(2016)은 초기 프로젝트 단계의 정보를 사용하여 시공 과정에서의 전기 소비량을 추정하는 방법을 제안했다. Cheng과 & Wu (2009)는 건설 관리 문제를 해결하는 데 사용되는 기계 추론모델을 제안하기 위해 유전알고리즘과 SVMs를 결합했다. Kim과 Teizer (2014)는 가설물 자동화 설계 방법을 제안했다.

3.5 운영

3.5.1 건물성능 모니터링

사물인터넷(IoT) 기술의 출현은 이런 류의 평가물들을 확장할 기회를 제공한다. 저비용 센서, 무선 네트워크의 사용성, 모바일 장치의 확대, 오픈 데이터 표준의 활용, 임베디드 디지털 기술의 통합은 데이터 수집의 복잡성과 비용을 대폭 줄였다. 실제로 이미 많은 데이터들이 수집되고 다른 맥락(예: 건물 관리 시스템(Building Management Systems, BMS)이나 건물 자동화 시스템(Building Automation Systems, BAS)에서 활용하기 위해 수집되고 있으나 아직 건물 설계자들에 의해 사용되지 못했다. 이러한 연속으로 작동하는 센서들은 현재의 한정된 어플리케이션들 보다 훨씬 더 유용한 데이터 스트림을 생성한다. (Dibley 외, 2011).

예를 들면, Peña 등(2016)은 건물 내 모든 센서에서 수집한 데이터를 사용하여 스마트 빌딩에서 에너지 효율의 이상을 감지하는 규칙 기반 시스템을 제안했다. Capozzoli 등(2015)은 사무소 건물의 오류 탐색을 위한 데이터 마이닝 시스템을 제안했다. Kim 등(2011)은 데이터 마이닝 방식을 통해 에너지 효율을 고려한 건축 설계물의 분석법을 보여주었다. Reffat와 Gero (2005)는 건물 관리를 위한 의사결정을 지원하기 위한 운영 데이터 채굴을 위해 객체 기반 CAD 시스템 및 3D 가상 환경을 데이터베이스 관리 시스템과 통합된 가상현실을 제안했다.

3.5.2 점유 모델링

대개의 건물 모니터링 시스템은 기반시설, 건물 시스템, 에너지 성능에 초점을 두고 사람들이 공간과 어떻게 상호 작용하는지에 대한 분석에 대한 관심은 매우 적다. 하지만 점유 패턴의 이해는 건물 성능을 예측하는 데 중요하다. Birt & Newsham (2009)은 설계자들이 일반적으로 건물 사용자(점유자)의 행태를 어떻게 잘못 예측하는 지에 대해 보여주었다. 몇몇 연구자들은 건물 센서에서 수집된 빅데이터를 사용하여 점유 패턴을 검출하는 방법을 제안했다. 대부분 건물 에너지 성능에 관련되는데, 예를 들어 D' Oca & Hong (2015)은 에너지 시뮬레이션에 사용할 수 있는 점유 스케줄의 패턴을 얻기 위해 2년 간 수집된 건물 빅데이터 스트림을 사용했다. Edwards 등(2012)은 미래의 주택 소비를 예측하는 데 가장 효과적인 기술을 결정하기 위해 건물 측정 빅데이터에 적용된 7가지 기계 학습 알고리즘을 평가했다. Zhao 등(2014)은 사무기기 소비 전력 데이터를 활용하여 거주자의 "수동적" 행태를 학습하는 "간접적인" 실용 데이터 마이닝 방법을 개발했다. Liang 등(2016)은 사무소 건물의 점유 스케줄 예측 모델을 제안했다.

4. 논의

4.1 설계 문제의 이해

건물 설계시 고려해야 할 가장 중요한 요인들 중 하나는 향후 점유자들의 요구지만 이는 아이러니하게도 적은 정보가 설계자들에게 유용하다는 화제들 중 하나일 것이다. 대지, 기후, 법적 요구 사항, 시공비 또는 현재 건축 트렌드에 관한 데이터는 일반적으로 미래 사용자의 특성에 관한 데이터보다 쉽게 얻을 수 있어서 아주 정보가 매우 풍부하다. 설계 문제를 정의하는 요구조건을 파악하기 위한 데이터의 주요 출처는 클라이언트 자신의 요구조건들에 대한 자기 인식인데 이는 일반적으로 객관적

이고 정확하거나 충분한 출처가 못된다.

뿐만 아니라 선행 연구는 클라이언트와 설계자 모두 자신의 가정들이 일반적으로 틀리다는 것을 보여주었다. 가장 일반적으로 설계자는 전문 지식과 직관력을 최대한 발휘하여 각 사례별로 "조정"되거나 "맞춤"화된 일반 데이터를 사용해야 한다. 실 크기에서 쾌적 수준 (예 : ASHRAE 표준)에 이르기까지 건물 설계는 "평균 사용자"의 설명에 크게 의존한다. 주된 가치가 맞춤화된 작업 산물인 산업의 경우에 이는 모순이 된다. 빅데이터 기술은 미래의 거주자들과 그들의 습관 및 선호도에 대한 풍부한 데이터를 제공할 수 있어서 설계자는 비교할 수 없는 세부 수준으로 그들의 요구조건들을 이해할 수 있다. 실제로 개인들의 세분화된 특성화는 빅데이터의 가장 보편적인 용도 중 하나이다. AECO 산업에서 동일한 방법을 사용할 수 있다고 사료된다. 예를 들어 소셜 네트워크와 지리적 위치 정보 기반의 장치 데이터를 이용하여 이동과 항구성의 행동 패턴을 파악할 수 있다. 건물 내장 센서 데이터를 사용하여 공간 이용 우선권을 추출할 수 있다.

상업용 건물은 고객 선호도 및 소비자 행동을 기반으로 설계될 수 있다. 학교는 학업 성취도와 밀접하게 연관된 공간 조건 분석을 기반으로 설계될 수 있다. 회의 공간은 다양한 사회 집단의 문화적 차이를 기반으로 설계될 수 있다. 특정 건물(예: 집)을 원하는 보통 사람조차도 단순한 자기 인식보다는 좀더 객관적으로 자신의 습관(선호도)을 알 수 있다. 그러나 사용자 특성화 과정이 어떻게 수행되어야하는지는 여전히 불투명하다. 데이터가 필요 이상으로 많고 설계물을 구축하는 데 가장 유용한 데이터가 무엇인지 모른다. 다른 산업 분야에서는 이미 어떤 원시 데이터가 가장 적합한 데이터인지 규명되었다. 정서 분석에서 실내 배치 시스템에 이르기까지 사용 가능한 개별 기술들이 존재하나 사용자의 공간 선호도를 이해하기 위해 이들을 단일 시스템으로 통합할 수 있는 프레임 워크나 방법이 없다. 현재의 애플리케이션들은 범위가 좁고 실용성이 낮다. 마지막으로 이러한 목적으로 사용되는 개인 데이터의 보안 및 프라이버시에 대한 법적 문제도 있다.

4.2 설계안 시뮬레이션

Lawson (1990)은 "대부분의 설계물들을 테스트하는 최상의 방법은 실제로 그들이 어떻게 잘 쓰이는지 기다리고 이해하는 것이다."라고 주장했다. 특정 엔지니어링 분야(예: 에너지나 구조 성능)에서는 정확한 시뮬레이션을 수행하는 시스템은 있지만 전체적인 건물 성능을 시뮬레이션할 수 있는 시스템은 거의 없다. 빅데이터는 새롭고 더욱 복잡한 시뮬레이션을 위한 기회를 열어준다. 설계 문제에 대한 정확한 정보를 얻을 수 있는 능력과 함께 설계안의 성능을 시뮬레이션할 수는 정보를 이용할 수 있는 능력이 있다. 오늘날 가장 강력한 응용 프로그램은 건물 분석과 인간 행동 분석을 결합한 것이다. 이는 화재시 대피 시뮬레이션과 같이 인간의 행태가 고도로 예측 가능한 상황에서 성공적으로 수행되었다. 수백 또는 수천 가지 사례의 데이터를 교차시켜서 어떤 공간 내에서 사람들의 움직임과 항상성에 대한 통계 모델을 생성할 수 있다. 시뮬레이션은 특정 사용자 집단에 맞춰 조정될 수 있고 사회, 문화적 특성이나 환경 조건에 따라 차별화된 공간 이용 패턴을 탐지할 수 있다. 설계자는 어떤 설계 특징이 공간 활용의 효율성을 극대화할 수 있는지 배울 수 있을 것이다. 이러한 시뮬레이션의 가치는 단독의 공간 계획 이상으로 확장되어 에너지 요구조건을 시뮬레이션하고, 구조적 과부하를 결정하고, 음향 조건을 예측하는데 사용된다. 일부 설계자는 설계에 대한 결정론적 접근 방식을 대표한다고 생각하기 때문에 이러한 유형의 시뮬레이션을 꺼린다. 건설 산업은 새로운 설계가 어떻게 수행될지에 대한 불확실성에 직면할 때 보수적이고 위험을 최소화하려는 경향을 보인다. 여전히 많은 연구 장애가 있는데 특히 어려운 점은 서로 다른 데이터 소스들을 시뮬레이션을 위한 통합 모델로 통합하기 위한 강력한 표준이 없다는 것이다. 실제로 크고 풍부한 데이터 세트를 다루는데 제한이 있지만, IFC (Industry Foundation Classes)와 Green Building XML (gbXML)은 주목할만한 발전을 이루었다. 빅데이터를 위해 견고한 표준을 개발하는 일은 AECO 업계 전반의 노력뿐 아니라 전자 및 IT 산업의 참여가 요구되기에 어려운 일이다. 더 중요한 것은 여러 유형의 프로젝트에 유용하고 의미 있는 개방형 시뮬레이션 프로세스를 작성하는 방법이 명확하지 않다는 것이다. 시뮬레이션의 질을 평가하는 측정기준과 프로토콜이 작성되어야 한다.

4.3 지속적인 설계 산물 평가

건축가들은 제안된 설계안이 어떻게 보이고 기능을 수행하는지에 대해 증명하기 위해 종종 "사실적" 이미지를 사용하지만 그러한 이미지와 예측을 실생활에서 효과적으로 시험하는 일은 일반적이지 않다. 거주 후 평가와 같은 현재의 건물 평가 방법은 주로 공학적 성능 지표에 한정된다. 건물 성능을 보다 종합적으로 분석하기 위한 다른 시스템들은 실현되기에 너무 복잡하다. 빅데이터 기술로 이러한 평가들은 기존의 에너지, 구조 또는 운영비와 관련된 내용을 초월할 수 있다. 예를 들어 클라이언트들은 공간 이용 효율, 사용자 쾌적도, 사회적 상호 작용이나 정서적 편안함과 같은 매우 복잡한 측정 항목을 평가할 수 있다. 사무소 건물의 근로자들을 위해 온 종일 개인의 쾌적 수준을 추적하는 시스템과 사용자의 손목 밴드에서 얻은 데이터로 체육관의 공기 조화 시스템의 성능을 평가한다. 실제로 건물을 평가하는 데 빅데이터 기술을 사용하는 주요 이유는 다른 목적으로 건물에 이미 내장된 전자 장치가 제공하는 데이터의 가용성이 증가하고 있다는 것이다. 사용가능한 광범위한 데이터는 전통적 방법으로는 실현 불가능한 새롭고 광범위하고 깊은 평가를 재고할 수 있는 기회다. 건물 외관의 매력도나 방의 아늑함과 같이 전통적으로 주관적이라고 여겨지던 개념은 수많은 셀카 사진으로부터 얻어진 관찰 가능한 지표를 기반으로 운용될 수 있다. 건물의 전자 장치의 가용성이 꾸준히 높아짐에 따라 지속적이고 영구적인 건물 모니터링 및 평가가 가능해졌다. 건물은 정적인 "완성된 프로젝트"가 아니라 끊임없이 진화하고 시간에 따라 변화하는 동적 객체로서 데이터에 대한 정보를 제공한다. 설계자들은 체계적이고 편향되지 않은 방식으로 건물로부터 배우고 클라이언트와 거주자들은 지속적으로 변화하는 요구에 더 잘 맞는 설계안과 공간에서 혜택을 받을 수 있다.

5. 결론

본고에서는 건물설계 분야에서 빅데이터 기술의 현황과 향후 방향을 검토했다. 조사 내용은 다양한 어플리케이션들이 사전 설계에서부터 건물 운영에 이르기까지 설계 프로세스의 전체 작업 영역을 아우르고 있음을 보여준다. 에너지 분석이나 구조 설계와 같이 전통적으로 데이터 중심으로 이해되는 작업 속에서 오늘날 가장 큰 진보를 이룬 수많은 사례들이 존재한다. 하지만 데이터에 의존하지 않고 직관이나 암묵적인 전문 지식에 의존하는 활동은 빅데이터 활용을 위한 더 흥미롭고 파격적인 기회를 제공한다. 빅데이터 분석의 암묵적 약속은 사실상 일반 데이터에 대한 의존성과 건물 설계를 구별하는 일화(逸話)적인 증거를 분리할 수 있는 탁월한 특수성, 객관성 및 통찰력으로 설계내용을 알려준다는 것이다. 빅데이터는 직감과 전문 지식을 배제하지 않고 알려준다. 더 많은 연구가 시급히 필요한 세 가지 응용 분야는 1) 클라이언트, 점유자, 대지, 기타 설계 조건들의 특수 데이터를 채굴하여 설계 문제를 이해하고 정의하는 것; 2) 기계 학습 방법으로 얻어진 통계 모델을 사용하여 설명적인 물리 모델 없이 복잡한 설계안들의 전체 성능을 시뮬레이션하고 예측하는 것; 3) 고도의 데이터 수집 및 분석 방법을 활용하는 새로운 건물 측정 기준들을 가지고 전생애주기 동안 설계 산물을 통합적이고 지속적으로 평가하는 것이다.

<원문제목> BIG DATA IN BUILDING DESIGN: A REVIEW

<원문출처> Journal of Information Technology in Construction