

실제 택배 데이터 기반으로 최적의 물류 허브 위치 도출

송하윤*, 한인수*

*홍익대학교 컴퓨터공학과

e-mail: hayoon@hongik.ac.kr

e-mail: haninsoo9989@gmail.com

Locating Logistics Hub with Parcel Delivery Data

Ha Yoon Song*, In Soo Han*

*Dept of Computer Engineering, Hong-ik University

요 약

택배 배달 서비스는 전국적 물류의 중요 부분이다. 노선이 잘못되면 시간과 비용 면에서 경제적 불이익이 발생할 수 있다. 이러한 위험을 최소화하기 위해 물류 회사나 IoT 기기 등에서 실제 배송 데이터 10만 개를 수집하여 최소 거리와 최소 시간의 관점에서 최적의 허브 후보 위치를 분석하였다. 택배의 공통 경로 데이터는 실제 택배의 경로 데이터에 Longest Common Route Sequence 알고리즘을 사용하여 계산하였다. 계산한 택배의 공통 경로 데이터를 기반으로, 최적의 허브 후보 위치를 설정하고, 이로부터 최적의 거리와 시간을 계산할 수 있다.

1. 서론

본 논문은 택배 배달 서비스의 경제적 불이익을 최소화하고, 최소 시간, 최소 거리의 관점에서 최적의 허브 후보 위치를 분석하는 것을 목표로 한다. 이를 위해, 실제 데이터를 수집하였다. 물류 회사나 IoT 기기를 활용하여, 10만 건의 배달 데이터를 수집하였다. 이 중 실제 유의미한 택배 데이터 2만 건을 선정 및 전처리하였다. 이 논문은 다음과 같이 구성되었다. 제2절에서는 우리가 사용한 LCRS 알고리즘을 포함하여 이 연구의 프로세스 개요에 대해 논한다. 제3절에서는 연구에 의해 분석한 최종적인 데이터 셋의 시각자료를 통해, 최적의 후보 위치를 선정한다. 제4절에서는 연구의 결과를 설명하며, 최적의 허브 후보 위치의 최적의 거리와 시간을 계산한다. 이를 통해, 결론을 내린다.

Beijing Jiaotong University의 Yunhe Ma는 가상의 데이터 셋을 기반으로 수식적인 접근을 함으로써, 최적의 허브를 도출하기 위한 연구를 하였다.[1] 본 논문은 실질적인 데이터를 기반으로, 최적의 허브를 제안하였을 때 얻을 수 있는 경제적 효익을 계산하고자 한다.

또한, National Taiwan University의 Yon-Chun Chou는 workload와 sojourn time을 줄이는 것을 목표로 유동 기간 동안 hub 운송량을 분석한다.[2] 본 논문은 최소 시간, 최소 거리 관점에서 최적의 물류 허브 위치를 제안하는 것을 목적으로 하고 있다.

2. 프로세스 개요

첫 번째 프로세스는 10만 건의 배달 데이터를 수집한다.

초기 데이터 셋은 [송장번호, 제품 카테고리, 택배의 방문지]로 이루어져 있다. 즉, 고객이 주문한 택배의 이동 경로를 5개의 경유지를 통해 제공한다. 그림 1은 원천 데이터 셋의 경로를 시각화한 그림이다.

두 번째 프로세스는 5개로 간략히 제공되어 있는 택배의 이동경로에 대해서 상세한 실제 경로 데이터를 추출한다. 즉, 고객이 주문한 택배의 이동 경로를 상세히 수집한다. 그림 2는 택배가 지나가는 실제 경로 데이터를 위경도를 이용하여 위치 데이터로 출력한 그림이다. 원형의 색상이 진할수록, 중복 방문 횟수가 많음을 의미한다.

세 번째 프로세스는, 수집한 택배의 이동 경로들에 대해 LCRS 알고리즘을 적용한다. LCRS 알고리즘이란, LCS 알고리즘을 본 연구에 알맞게 발전시킨 알고리즘이다. LCS 알고리즘은 두 문자열이 있을 때, 서로 공통으로 가지고 있는 문자열의 길이를 구하는 알고리즘이다.[3] 본 논문은 이를 발전시켜, 두 경로 배열이 있을 때, 서로 공통으로 방문하는 위치 데이터들을 구하는 LCRS 알고리즘을 구축하였다. 그림 3은 LCRS 알고리즘을 통해 추출한 공통 방문 위치 데이터를 시각화한 그림이다. 원형의 색상이 진할수록, 중복 방문 횟수가 많음을 의미한다.

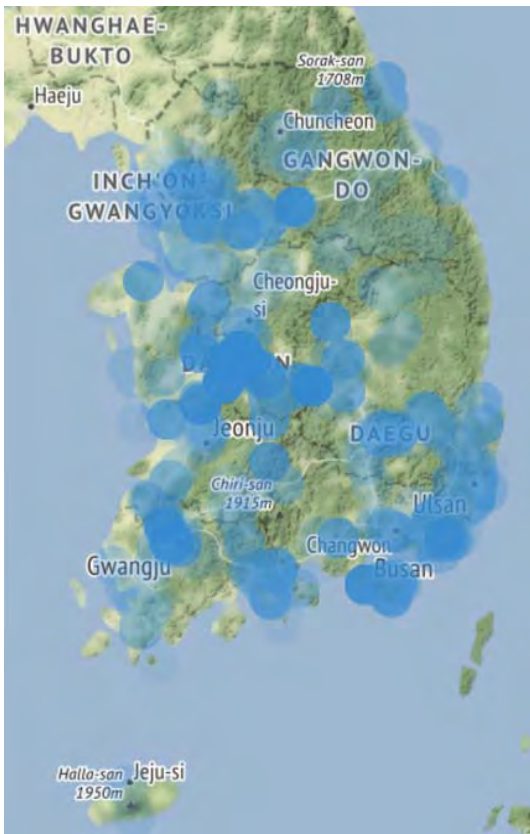
네 번째 프로세스는, 수집한 LCRS 데이터 셋에 대해 각 location에 택배가 몇 번 방문했는지를 Counting 하게 된다. 이를 통해, 공통 방문 경로 중 가장 많이 방문하게 되는 location을 Voting 할 수 있다. 원형의 크기가 클수록 최종 방문 횟수가 많음을 의미한다.

마지막으로 다섯 번째 프로세스는, Voting 결과를 기반으로, 가장 경제적인 Hub location을 candidate로 선정하

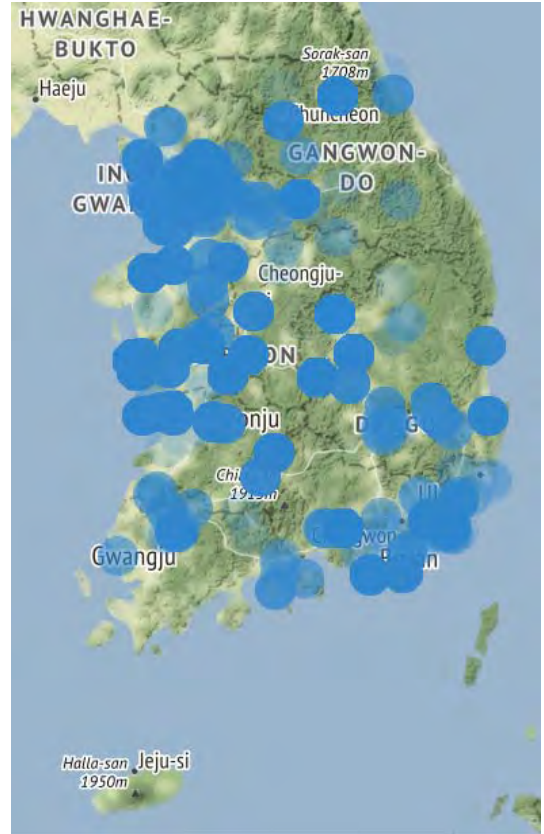
고, 각 candidate hub에 대해 최소 거리, 최소 시간 관점의 cost를 계산한다. 이를 통해 가장 경제적이고 불이익을 최소화할 수 있는 Sub Hub 위치를 제안한다.



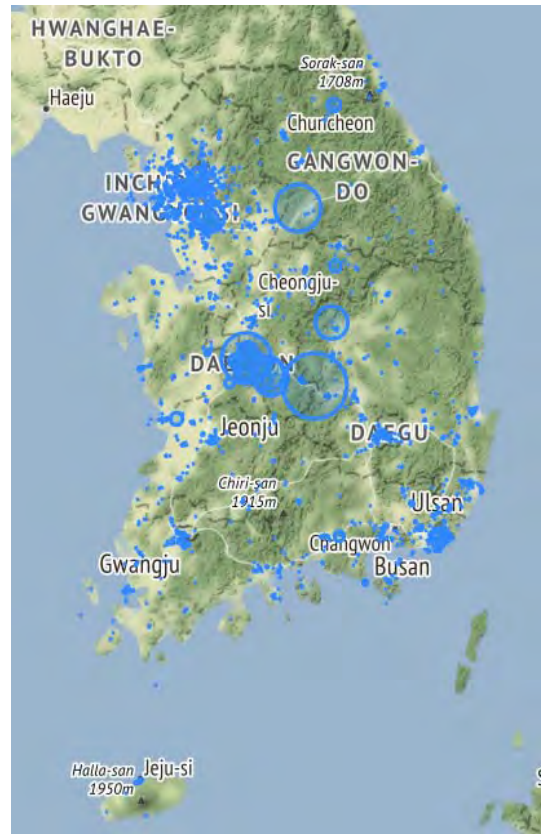
(그림 1) 원천 데이터 셋 경로 시각화



(그림 2) 경로 데이터 셋 시각화



(그림 3) LCRS 알고리즘 결과 시각화



(그림 4) Voting 결과 시각화

3. 실험 결과

그림 4는 Voting 결과, 어느 위치 데이터에 방문 횟수가 많은지를 시각화한 그림이다. 그림 4의 파란색 원이 큰 지점이 Counting이 많은 지점이다. 파란색 원이 작은 지점은 Counting이 적은 지점이다. 이를 통해, 서울에는 공통 방문 횟수가 큰 location는 적지만, 매우 다양한 지점에 물류량이 존재하는 것을 알 수 있다. 즉, 서울이 유통량이 가장 많은 지점임을 알 수 있다. 두 번째로 옥천 허브에 해당하는 지점이 다양한 지점에 물류량이 존재하는 것을 알 수 있다. 세 번째로 부산 지역에 물류량이 많은 것을 그림 4를 통해 알 수 있다.

또한, 큰 원이 경부 고속도로에 해당하는 지역에 존재하는 것을 확인할 수 있다. 이를 통해, 서울과 부산 지점에서 택배의 요구량이 가장 많고, 서울과 부산 사이에 해당하는 대한민국 최대의 물류 허브인 옥천 Hub에서 유통 및 공통 방문 횟수가 가장 많은 것을 확인할 수 있다.

즉, 서울과 부산 사이의 경로에 해당하며, 옥천 Hub만을 이용했을 때 보다, 최소 거리, 최소 시간 관점에서 경제적 불이익을 해소할 수 있는 지점을 옥천 Hub 근방의 지역에서 제안할 수 있다.

본 논문은 “보은”, “문경”, “금산” 세 지역을 후보 허브로 선정하였다. 그림 5는 선정된 후보 허브와 현재 최대 허브인 옥천 허브의 위치이다. 그림 5에서 A가 옥천 허브, B가 금산, C가 보은, D가 문경이다.



(그림 5) 후보 허브

4. 연구 결과 및 연구 전망

다섯 번째 프로세스에서는 선정된 후보 허브 “보은”, “문경”, “금산”에 대해서 최소 거리, 최소 시간의 관점으로 기존의 대형 물류 허브인 “옥천” 허브와 비교한다. 옥천 허브를 지나는 대신, 후보 허브를 지나는 경우에 소요되는 시간 및 거리를 옥천 허브와 비교해보았다. 그 결과, 보은과 문경에 허브를 설치하는 경우, 옥천보다 비효율적이라는 결과가 나왔다. 그에 비해, 금산에 허브를 설치하는 경우, 택배 1건 당 평균 2.5km의 거리적 이익이 발생한다는 결과가 나왔다. 이는, 옥천 대신 금산에 허브를 설치하는 경우, 본 논문에서 수집한 데이터 10만 개의 데이터 셋에 대해서, 평균 25,000km의 거리적 이익을 얻을 수 있다는 것을 의미한다.

본 논문은 데이터를 수집하여 사용하기에, Outlier가 많을 수 있다. 하지만 물류 회사로부터 전체 데이터를 수령하여 분석한다면 더 정확한 결과를 도출할 수 있을 것으로 예상된다. 그리고 오스트리아, 베트남 등 한국과 비슷한 환경을 갖고 있지만, 물류 산업이 활성화되지 않은 해외를 타깃으로 연구를 진행한다면, 최적의 허브 위치를 제안할 수 있을 것으로 기대된다. 추가적으로 데이터 셋이 매우 크므로, GPGPU를 사용하여 빅 데이터를 효과적이고 빠르게 분석할 수 있다.

Acknowledgement

이 연구는 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행되었다. (NRF-2017R1D1A1B03029788)

참고문헌

[1] Jijun Pang, Yunhe Ma, Xifu Wang. Research on hub location and routing distribution for hub-and-spoke logistics network. 2016 6th International Conference on Machinery, Materials, Environment, Biotechnology and Computer, January 2016.

[2] Hui-Min Chen Yon-Chun Chou, Yao-Hung Chen. Pickup and delivery routing with hub transshipment across flexible time periods for improving dual objectives on workload and waiting time. Transportation Research Part E: Logistics and Transportation Review, 2014.

[3] Harri Hakonen Bergroth, Lasse and Timo Raita. A survey of longest common subsequence algorithm. SPIRE 2000(Seventh International Symposium on String Processing and Information Retrieval, IEEE), 06 August 2002.