

유전자 알고리즘을 이용한 음식 배달 최적화 기법

양소연¹, 임유진²

¹ 숙명여자대학교 빅데이터분석융합학(협동과정)

² 숙명여자대학교 IT 공학전공

syayang@sookmyung.ac.kr, yujin91@sookmyung.ac.kr

Optimization with Genetic Algorithms for Food Delivery Dispatch

Soyeon Yang¹, Yujin Lim²

¹Dept. of Big Data Analysis Convergence, Sookmyung Women's University

²Dept. of IT Engineering, Sookmyung Women's University

요 약

과열되는 배달 시장의 경쟁 속에서 수많은 배달원이 속도전에 내몰리고 있다. 배달 앱 시장에서는 속도전의 승리를 위해 단건 배달 서비스를 내놓았지만 이러한 배달 경쟁은 배달비 인상으로 이어져 소비자의 부담으로 돌아왔다. 본 논문에서는 배달 업무의 고른 분배를 통해 배달원들의 경쟁을 완화하고 전체 배달 시스템의 처리량과 신뢰도를 향상하고자 하였다. 따라서 목적함수의 최적화와 무작위성이라는 특징을 가진 유전 알고리즘을 활용하여 배달원들 간 배달 업무의 고른 분배를 통해 시스템 성능을 향상시켰다. 실험을 통해 기존의 배치 기법에 비하여 제안 알고리즘에서의 성능이 향상되었음을 확인하였다.

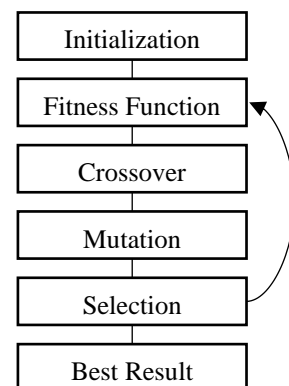
1. 서론

통계청이 발간한 “2022년 2월 온라인쇼핑 동향”에 따르면 온라인 주문 배달음식 거래액은 코로나 19가 확산한 2020년 대비 약 90% 이상 증가하였다[1]. 배달 앱 시장이 커지면서 경쟁이 심화하였고 지나치게 큰 중계 수수료 부담 문제와 단건 배달의 속도 경쟁 등의 논란은 끊임없이 사회적 문제로 제기되고 있다. 이러한 배달 대행업체들의 경쟁은 배달비 상승 등을 일으킴으로써 그 부담이 소비자에게 전가되고 있다. 본 논문은 최적화된 배달원 배치를 통한 배달 처리량의 증가뿐만 아니라 배달 신뢰도 향상을 통하여 배달 대행 업체들 간의 과열된 경쟁을 완화하고자 하였다. 이를 위하여 본 논문에서는 유전 알고리즘의 무작위성을 활용하여 배달원들에게 배달 업무를 고르게 분배하는 알고리즘을 제안한다. 배달 업무의 고른 분배를 통해 배달원들의 일정한 수입을 보장함과 동시에 속도 경쟁으로 인한 인명사고의 위험 및 업무의 과중을 줄임으로써 효율적인 업무 수행을 할 수 있도록 지원하고자 한다.

유전 알고리즘은 다윈의 적자생존 이론을 바탕으로 개발된 최적화 모델이다[2]. 유전 알고리즘은 외판원 문제(TSP), 스케줄링 문제 등 NP-Hard에 속하는 문제들을 해결하는데 응용되고 있다.

본 논문에서는 배달원들 간 배달 업무를 고르게 분배하기 위하여 유전 알고리즘의 적합도 함수(fitness function)를 배달원의 배달 건수의 표준편차를 고려하여 제안한다. 또한, 제안하는 알고리즘이 단순히 식당과 가장 가까운 거리의 라이더를 배차하는 것과 비교하였을 때 어떠한 성능차이를 보이는지 비교하였다.

2. 시스템 모델



(그림 1) 유전 알고리즘의 구조

본 논문에서는 미국의 대형 음식 배달 대행업체인

“그럽허브(Grubhub)”가 깃허브(Github)에 공개한 배달 주문 데이터셋을 활용하였다[3]. 해당 데이터셋은 116 개의 음식점 정보 R 과 505 개의 주문 정보 O , 그리고 113 개의 배달원 정보 C 로 구성되어 있다. 각 음식점 $r \in R$ 은 해당 음식점의 위치 좌표 (r_x, r_y) 를 가지고 있다. 주문 정보 $o \in O$ 에는 음식 주문과 관련하여 음식 주문 시각 a_o , 음식 준비 완료 시각 e_o , 그리고 배달지 위치정보 (o_x, o_y) 가 있다. 배달원 $c \in C$ 는 출근 시각 a_c 과 퇴근 시각인 e_c 뿐만 아니라 시간 t 에서 배달원의 위치인 (c_x^t, c_y^t) 가 있다.

본 논문에서 제안하는 유전 알고리즘의 구조는 (그림 1)과 같으며 데이터셋을 기반으로 주문 o 이 발생한다고 가정한다. 생성되는 주문 건에 따라 음식 주문시간이 배달원의 근무시간 $e_c - a_c$ 에 해당하는 배달원의 정보를 유전 인자로 설정하였다. 유전자는 배달원별 기 배달 건수 d_c^t 와 배달원의 현재 위치 (c_x^t, c_y^t) , 배달원의 근무 시작시각 a_c 와 근무 종료시각 e_c , 그리고 배달원의 신뢰도 f_c 로 총 6 개로 구성되며 염색체(chromosome)의 수는 시간 t 에서 근무 중인 C_t 명의 배달원의 수와 같게 설정하였다. 배달원당 하루 배달 건수 d_c 는 배달 라이더의 하루 평균 배달 건수가 약 29.7 건[4]이라는 조사 결과에 따라 최대 30 건 이내에서 무작위로 부여하였다. 신뢰도 f 는 0.95 에서 1 사이의 값으로 랜덤하게 설정하였다.

제안하는 알고리즘에서는 배달 업무의 고른 분배를 위해 적합도 함수에서 각 배달원 간의 배달 건수 표준편차를 고려하였고 이것은 식 (1)과 같다. \bar{d}^t 는 d_c^t 집단의 평균값이다.

$$S_d = \sqrt{\frac{\sum_{c=1}^{C_t} (d_c^t - \bar{d}^t)^2}{C_t}} \quad (1)$$

또한, 배달원 c 가 근무시간 동안 이행한 배달들이 예상 시간 내에 배달 완료되었는지를 나타내는 신뢰도를 식 (2)와 같이 정의하였다.

$$F_c = \frac{d_c}{e_c - a_c} \quad (2)$$

구해진 표준편차 S_d 와 신뢰도 F 을 사용하여 적합도 함수를 식 (3)과 같이 계산하였다.

$$Fitness = \alpha \frac{1}{S_d} + (1 - \alpha) \cdot F_c(3)$$

배달 건수의 고른 분배를 위하여 배달 건수에 대한 표준편차 S_d 는 작을수록 이와 달리 신뢰도 F 는 클수록 적합도 값이 커지도록 설정하였다.

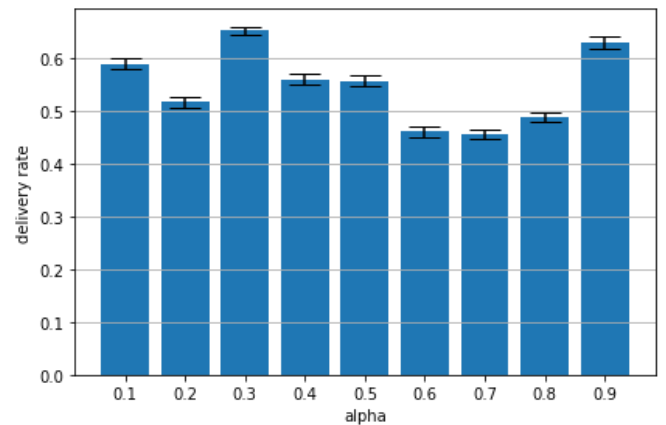
3. 제안하는 알고리즘

본 논문에서는 유전 알고리즘을 활용하여 배달원의 고른 업무 배분 및 배달 신뢰도 향상을 위한 알고리즘을 제안한다. 해당 알고리즘에서는 공개 데이터셋을 기반으로 주문을 발생시키고 주문이 들어온 때에 근무하는 라이더의 정보를 수집한다. 수집된 라이더의 정보는 유전 알고리즘의 초깃값이 되어 적합도 함수를 거친다. 적합도 함수에 따라 각 유전자는 표준편차 S_d 는 작게 만들고 신뢰도는 증가시키는 Fitness 값을 생성한다. 해당 값은 부모 인자가 되어 교차와 변이를 거쳐 종료 조건이 될 때까지 진화과정을 반복한다.

4. 실험 및 성능 비교 분석

해당 논문에서는 제안한 알고리즘의 성능을 측정하기 위해 배달 주문이 급증하는 시간을 대상으로 하였다. 유전 알고리즘에서는 파라미터 값인 인구수(population)와 세대수(generation)에 따라 훈련 횟수가 정해진다. 본 실험은 인구수를 500, 세대수를 1000으로 설정하여 진행되었다.

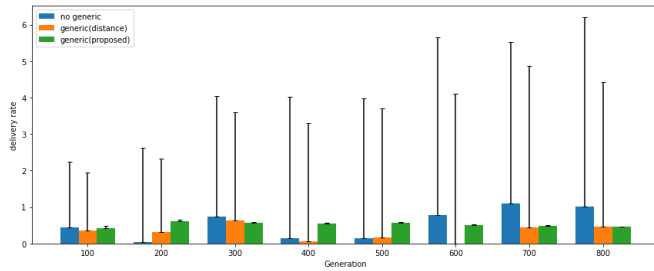
적합도 함수의 가중치인 α 값에 0.1부터 0.9까지의 변화를 주며 어떤 값에서 가장 낮은 표준편차 S_d 가 나왔는지 즉, 어떤 값에서 가장 배달 업무의 배분이 잘 되었는지 살펴보았다. 그 결과 α 가 0.3일 때 배달원 간의 배달 건수 표준편차가 값이 가장 작았으며 α 값의 변화에 따른 배달 건수의 분배 결과는 그림 2와 같다.



(그림 2) 가중치 α 값 변화에 따른 표준편차

이러한 결과를 바탕으로 적합도 함수의 가중치 α 값을 0.3으로 설정하여 두 가지 경우와 비교하여 실험을 진행하였다. 첫 번째는 단순히 배달원과 식당 간의 거리만 고려하여 식당과 가까운 배달원을 배치한 경우(no generic)이며 두 번째는 유전 알고리즘의 적합도 함수를 이용하여 식당과 가까운 배달원을 배

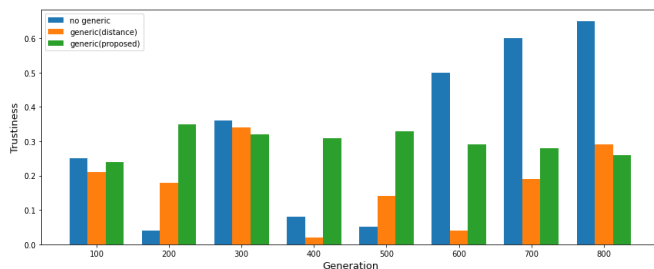
치한 경우($\text{generic}(\text{distance})$)이다. 비교 결과는 그림 3 과 같다.



(그림 3) 세대수에 따른 표준편차

배달 건수 면에서 no generic 기법은 평균 0.54 건을 배달하였고 $\text{generic}(\text{distance})$ 기법은 평균 0.30 건을, 그리고 제안 알고리즘인 $\text{generic}(\text{proposed})$ 기법은 평균 0.51 건을 배달한 것으로 나타났다. 제안 알고리즘과 no generic 기법이 서로 비슷하게 좋은 결과를 보였고 표준편차는 세 모델 중 제안 알고리즘이 가장 낮았다. 이는 제안 알고리즘이 배달 업무 분배를 고르게 잘하였다고 볼 수 있다. 상대적으로 no generic 과 $\text{generic}(\text{distance})$ 기법은 배달 업무의 분배가 효율적으로 이루어지지 않았고 이러한 문제를 제안 알고리즘을 통해 해결할 수 있음을 보여준다.

신뢰도 면에서 no generic 기법과 제안 알고리즘의 신뢰도가 약 0.3 으로 비슷하게 측정되었고 $\text{generic}(\text{distance})$ 기법은 평균 0.17 로 가장 낮았다. 해당 실험 결과는 그림 4 와 같다.



(그림 4) 세대수에 따른 신뢰도

또한 no generic 과 $\text{generic}(\text{distance})$ 기법은 세대수에 따라 신뢰도의 변화 폭이 컸지만, 제안 알고리즘의 신뢰도는 비교적 변화 폭이 작은 것으로 확인하였다.

5. 결론

제안하는 모델에서는 배달 대행업체들 간의 과열된 경쟁을 완화하고자 배달원간 배달 건수의 표준편차를 줄이면서도 배달 신뢰도를 향상하고자 하였다. 이 문제를 해결하기 위해 유전 알고리즘을 활용하였고 적합도 함수에서 배달원 간의 배달 건수 표준편차를 고려하였다. 그 결과 단순히 가까운 거리의 배달원에게

업무를 부여하는 것보다 해당 알고리즘을 사용하였을 때 고른 업무 배분이 이루어졌다는 것을 확인하였다. 또한 신뢰도 면에서도 제안 기법이 상대적으로 일정한 신뢰도를 유지하고 있음을 확인하였다.

향후 연구에서는 배달업무의 고른 분배뿐만 아니라 배달원의 신뢰도를 향상시킴과 동시에 배달원이 최적의 배달 경로를 찾을 수 있도록 하는 연구를 수행하고자 한다.

사사문구

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 ICT 혁신인재 4.0 사업과 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (IITP-2022(2022)-RS-2022-00156299, No. 2021R1F1A1047113)

참고문헌

- [1] 통계청, “2022 년 2 월 온라인쇼핑동향”, 2022
- [2] D. Whitley, "A Genetic Algorithm Tutorial." Statistics and Computing, vol. 4 no.2. pp.65-85, 1994.
- [3] D. Reyes, A. Erera, M. Savelsbergh, S. Sahasrabudhe and Ryan J. O'Neil, "The Meal Delivery Routing Problem," Optimization Online, vol. 6571, Apr. 2018
- [4] 이고은, “시간당 배달건수 제한은 라이더 사고 막을 수 있을까,” 매일노동뉴스, 2022.01.27.