

## [Review]

본 논문은 기존의 Contextual Multi-armed bandit의 보상 함수가 시간이 지남에 따라 User의 선호도의 변화를 반영할 수 없다는 한계를 해결하기 위해서 Particle learning을 기반으로 한 dynamical context drift model을 제안하였다.

제안된 모델은 random walk particles set으로 모델링 되고, fitted particle이 reward function을 학습한다. Particle learning을 통해 모델이 context의 변화를 포착할 수 있으며 context의 latent parameters를 학습할 수 있다.

모델의 목표는 latent parameters variables와 latent state random variables를 sequential observed data를 통해 추론하는 것이다. 하지만 latent state random variable은 random walk 방법에 의존하기 때문에, 저자는 latent parameters, latent state random의 distribution을 학습하기 위해 Sequential Monte Carlo sampling과 Particle learning을 이용하여 random walk 방법을 사용하였다.

본 논문에서 제안된 모델은 기존에 존재하는 Bandit algorithms를 개선하여 보다 contextual dynamics한 상황 tracking을 가능하게 하였으며 CTR에 대하여 personalized recommendation performance를 향상시켰다

## [본문 URL]

<https://velog.io/@tobigs-recsys/Paper-Review-2016-Chunqiu-Zeng-Online-Context-Aware-Recommendation-with-Time-Varying-Multi-Armed-Bandit>