

Далее будет представлен поэтапный пайплайн модели генерации графов при помощи нетривиальных распутанных представлений. Пайплайн можно видеть на Рис. 1-2.

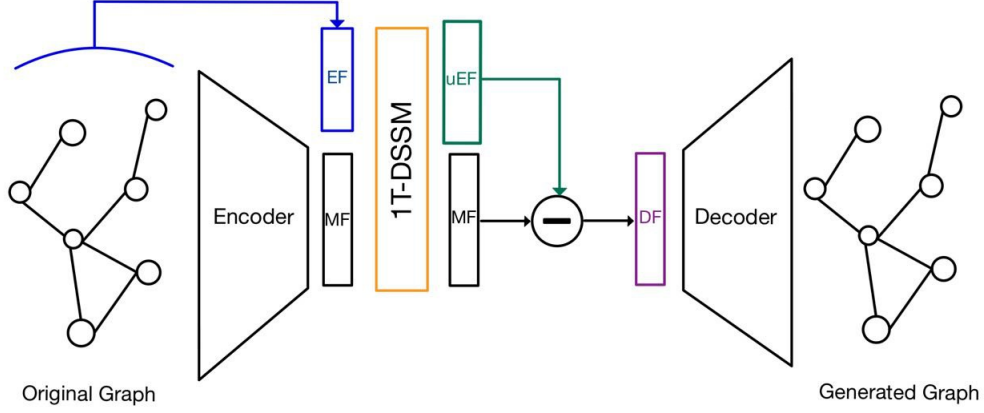


Рис. 1: Схема модели с генерацией только на основе сложных признаков.

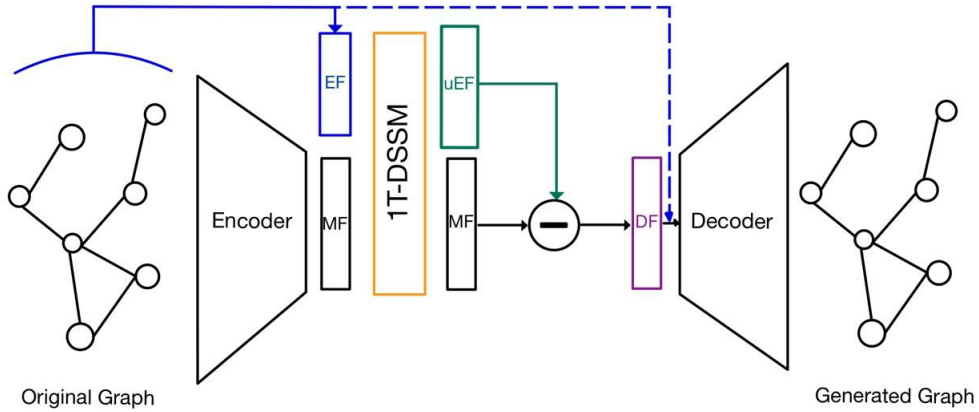


Рис. 2: Схема модели с генерацией на основе сложных и простых признаков.

Здесь EF – это представление простых признаков, которые можно достаточно быстро посчитать точными алгоритмами (число вершин, число ребер, максимальная степень и т.п.). DF – это представления сложных признаков графа (например максимальная смежная центральность). MF – это смешанное представление сложных и простых признаков графа. uEF – представление, которое построено на основе EF, и максимально приближено к MF.

Первым этапом является обучение распутанного представления на основе смешанных признаков графа. Это будет происходить при помощи beta-VAE (Рис. 3).

$$\begin{aligned}
 ELBO &= E_{q_\phi(z|x)} \log \left[\frac{p_\theta(z, x)}{q_\phi(z|x)} \right] = \\
 &= E_{q_\phi(z|x)} [\log p_\theta(x|z)] - KL(q_\phi(z|x) || p_\theta(z)) = \\
 &= \log p_\theta(x) - \beta \cdot KL(q_\phi(z|x) || p_\theta(z|x)) \leq \log p_\theta(x)
 \end{aligned}$$

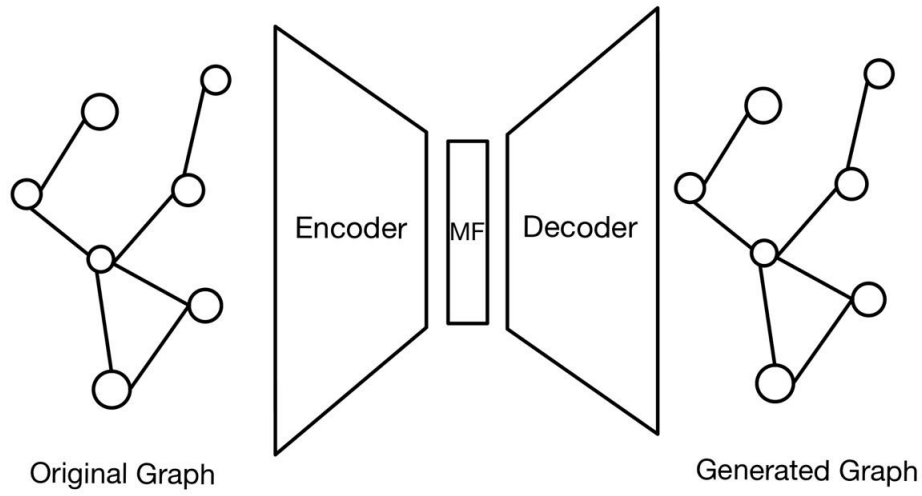


Рис. 3: β -GraphVAE.

Следующим этапом идет обучение модели 1T-DSSM (Рис. 4). Модель будет использовать обученные представления из прошлого этапа и будет стремиться используя только простые признаки, получить представление смешанных. Часть пайплайна обучения предлавлена на Рис. 5.

При обучении 1T-DSSM мы будем оптимизировать следующий функционал:

$$\text{sim}(C, D) = \frac{C^T \cdot \bar{D}}{\|C\| \cdot \|\bar{D}\|} = \frac{MF^T \cdot uEF}{\|MF\| \cdot \|uEF\|}$$

Благодаря использованию 1T-DSSM мы можем исключить из MF большую часть влияния простых признаков просто отняв от MF uEF и получить представление только сложных признаков графа DF.

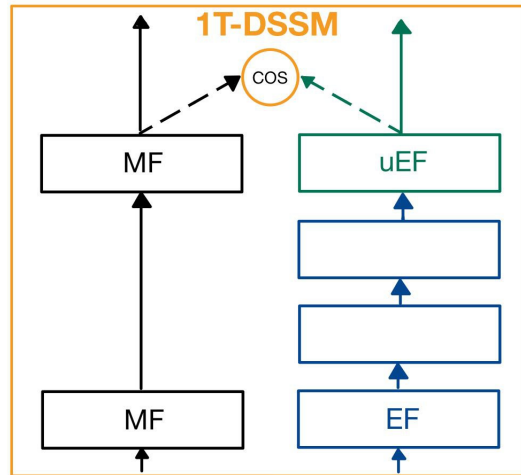


Рис. 4: Модель однобашенной DSSM.

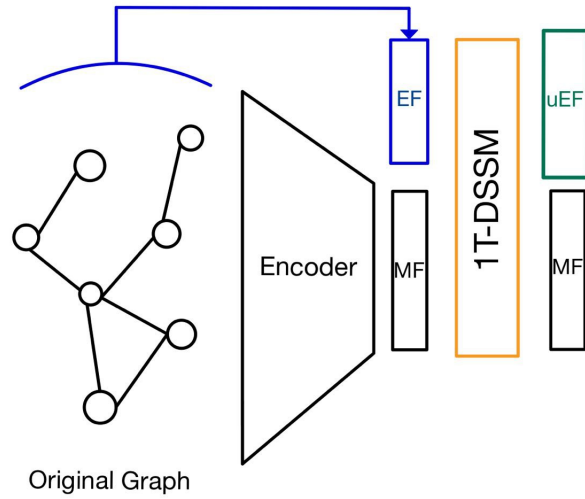


Рис. 5: Часть модели на которой будет обучаться модель 1T-DSSM.

После подготовительных этапов, весь пайплайн будет обучаться вместе в два этапа:

- Несколько эпох обучения VAE с фиксированными весами 1T-DSSM
- Несколько эпох обучения 1T-DSSM при фиксированных весах VAE

Итогом этого пайплайна будет возможность генерировать графы с фиксированными простыми признаками и случайными сложными, что позволит генерировать похожие графы с похожей макроструктурой, но сильно отличающейся микроструктурой.