

# Предсказание появления ребра в темпоральном графе

*Команда №2  
Волокитин Егор  
Камынин Павел  
20.Б12-ПУ*

# Стек технологий

- ЯП: C++20
- Компилятор: *gcc*
- Сборщик: *make*
- Библиотека базовых структур данных: *STL*
- ML: *mlpack*

# Оценка свойств графа

Граф хранился в структуре: `std::vector<std::unordered_set<int>>`

radoslaw_email		
Vertex count	167	
Edge count	3250	
Density	0.234	
Weak comp. count	1	
Main comp. to graph	1.000	
Main component radius	3	
Main component diameter	5	
Main component 90%	3	
Cl koef	0.296	
Assort koef	-0.295	

test_4		
Vertex count	62	
Edge count	159	
Density	0.084	
Weak comp. count	1	
Main comp. to graph	1.000	
Main component radius	5	
Main component diameter	8	
Main component 90%	5	
Cl koef	0.129	
Assort koef	-0.044	

team_2		
Vertex count	1098	
Edge count	60151	
Density	0.100	
Weak comp. count	1	
Main comp. to graph	1.000	
Main component radius	Rand: 2	
Main component radius	Snow: 2	
Main component diameter	Rand: 3	
Main component diameter	Snow: 3	
Main component 90%	Rand: 2	
Main component 90%	Snow: 2	
Cl koef	0.050	
Assort koef	0.002	

## Статические признаки для ребра (1, 2)

Структура данных признаков: `std::map<std::pair<int,int>, std::vector<double>>`

	radoslaw	team_2	reed	test_2	test_5	test_7	middlebury
<b>CN</b>	67	12	19	7	0	583	2
<b>AA</b>	16.8858	2.527	4.724	6.1307	0	255.859	0.3824
<b>JC</b>	0.5075	0.052	0.186	0.3888	0	0.07612	0.011
<b>PA</b>	8970	14224	3504	144	69	4.7*10^6	8262

# Темпоральный граф

Структура данных для хранения графа с временными пометками: `std::map<std::pair<int, int>, std::set<int>>`

Почему set? Потому что это ускорит вычисление статических признаков

*Common Neighbours (CN)* The CN feature is equal to the number of common neighbours of two nodes.

$$CN_{\text{static}}(u, v) = |\Gamma(u) \cap \Gamma(v)| \quad (1)$$

*Adamic-Adar (AA)* The AA feature considers all common neighbours, favouring nodes with low degrees (Adamic and Adar 2003).

$$AA_{\text{static}}(u, v) = \sum_{z \in \Gamma(u) \cap \Gamma(v)} \frac{1}{\log |\Gamma(z)|} \quad (2)$$

*Jaccard Coefficient (JC)* The JC feature is similar to the CN feature, but normalises for the number of unique neighbours of the two nodes.

$$JC_{\text{static}}(u, v) = \frac{|\Gamma(u) \cap \Gamma(v)|}{|\Gamma(u) \cup \Gamma(v)|} \quad (3)$$

*Preferential Attachment (PA)* The PA feature takes into account the observation that nodes with a high degree are more likely to make new links than nodes with a lower degree.

$$PA_{\text{static}}(u, v) = |\Gamma(u)| \cdot |\Gamma(v)| \quad (4)$$

# ROC AUC Score

	<b>radoslaw</b>	<b>bitcoin alpha</b>	<b>usocial</b>	<b>mathoverflow</b>
<b>Static</b>	85%	86%	75%	85%
<b>Temporal</b>	88%	90%	83%	91%

**Спасибо за внимание!**