

УДК 796.058.4

Подкорытов Глеб Игоревич,
студент,
Институт Машиностроения, Материалов и Транспорта,
Санкт-Петербургский Политехнический университет
г. Санкт-Петербург, Российская Федерация

ПРИМЕНЕНИЕ АНАЛИТИКИ В ФУТБОЛЕ

Аннотация:

В статье приведены результаты исследования соответствия аналитических данных с реальными результатами футбольных матчей. Рассмотрены примеры использования тех или иных данных, метрик для анализа игры и её результата.

Ключевые слова:

Аналитика, метрика, статистика, футбол.

Введение

Алгоритм начинает играть важнейшую роль во всех областях жизни человека [1], способствуя преобразованию общества в цифровую систему [2]. В спорте, в частности в футболе, стремительно внедряются и развиваются цифровые технологии [3]. Многие футбольные команды, выступающие в высших лигах, имеют аналитические штабы, которые собирают необходимую информацию, результаты исследования которой можно использовать для планирования тренировок и более тщательной подготовки к соперникам, матчам или даже ко всему сезону.

Методы исследования

В настоящее время аналитики используют огромное количество метрик, по которым можно собирать данные. Есть командные и индивидуальные показатели. Рассмотрим только три из них – xG, DEER, PPDA.

Основная идея xG заключается в придании каждому удару коэффициента опасности, который определяет колеблющуюся от 0 до 1 вероятность забить гол после удара. Распределение баллов зависит от точки, из которой нанесен удар, части тела, которой произведен удар, последнего паса, типа атаки. При помощи базы из сотен тысяч примеров выделяются для каждого удара свои коэффициенты, которые по ходу игры суммируются и в итоге виден «ожидаемый счет». Данная метрика позволяет судить о качестве голевых моментов команды и её реализации. [4,5].

DEER – количество успешных передач, сделанных в 18 – ти метровую зону (20 ярдов) от ворот соперника, исключая навесы (рис. 1), т.е. оцениваем по этой метрике созидательный компонент игры или как команда играет в позиционном нападении. Этот показатель во многих случаях коррелирует с результатом матча. [6,7]

PPDA (Passes allowed Per Defensive Action) – показатель интенсивности прессинга. Рассчитывается как отношение числа успешных передач соперника на количество всех активных оборонительных действий команды. Чем больше показатель интенсивности, тем хуже команда прессингует соперника, позволяя ему «спокойнее» владеть мячом. [7].

Мы проанализируем эти метрики и их связь с результатами матчей. Не всегда качество игры перерастает в результат из-за реализации, класса игроков и т.д. Но у каждой команды есть свои тактика, принципы игры и исполнители разного качества. Для наибольшей объективности и независимости исследования, в качестве объекта выступает Zenit СПб. Это чемпион последних двух сезонов, в котором тренер работает три года, и в составе происходят только точечные изменения, которые не меняют тактику. Мы возьмём матчи только одного

турнира – Российской Премьер-Лиги и проанализируем 113 матчей Зенита в РПЛ за 3 полных сезона (2017-2020 гг.) и 1 неполный сезон 2020-2021 (до 17.03.2021) по следующим критериям: 1) xG, 2) xG opp. (opp. – opponent - соперник), 3) DEEP, 4) DEEP opp., 5) PPDA, 6) PPDA opp., 7) итог матча и 8) Chances. Chances – это самостоятельно вычисляемая независимой аналитической компанией вероятность результата, исходя из статистики. В статье используются данные, выкладываемые в публичном виде аналитической компанией на сайте <https://understat.com> [8].

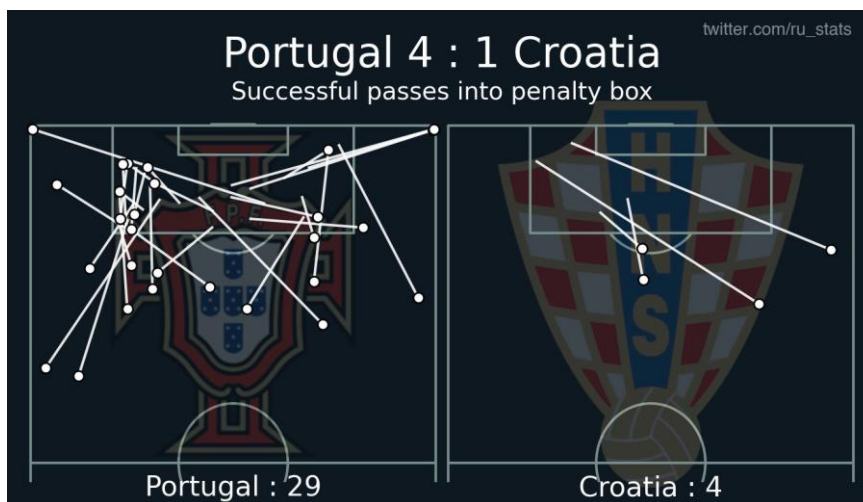


Рисунок 1 – Визуализация DEEP

Результаты исследования

В большинстве матчей статистика связана с результатом, причем команда может превосходить соперника по всем показателям, а может не превосходить в одной метрике, но доминировать в других. Первую категорию представляют матчи, когда команда имеет преимущество перед соперником по всем статистическим показателям - 34 из 113 (30,09%). Часто они заканчиваются разгромным счетом. Примеры: 30 тур 2017/2018 Зенит-СКА Хабаровск 6:0(рис.2); 17 тур 2020/2021 Зенит-Урал 5:1(рис.3); 10 тур 2019/2020 Зенит-Рубин 5:0. Как мы видим, вероятность (Chances) подтверждает закономерность результата. Но есть и примеры (12 матчей, 10,62%), когда игра является равной как по статистике, так и по результату. Например, матч 22 тура сезона 2017/2018 Ростов-Зенит, в котором команды не нашли победителя (табл. 1).

Zenit St. Petersburg	TEAMS	SKA-Khabarovsk
100%	CHANCES	
6	GOALS	0
4.19	xG	0.21
30	SHOTS	5
14	SHOTS ON TARGET	1
13	DEEP	3
6.41	PPDA	22.50
2.99	xPTS	0.00

Рисунок 2 – Зенит-СКА

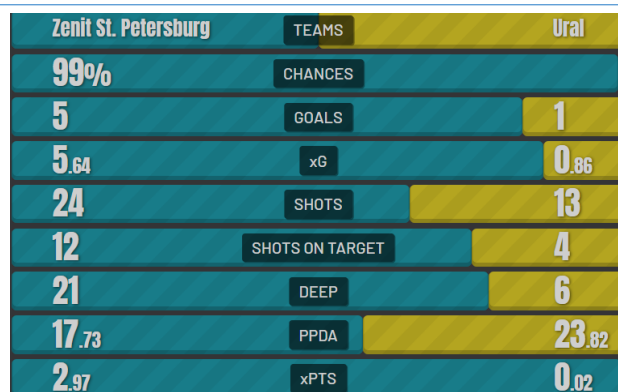


Рисунок 3 – Зенит-Урал

Таблица 1 – Ростов-Зенит

xG	xG opp.	DEEP	DEEP opp.	PPDA	PPDA opp.	Итог	Chances
0,3	0,57	3	5	10,9	9,83	0:0	49%

Но обнаружены и другие закономерности, когда победитель превосходил соперника всего по двум или даже по одной метрикам – 32 матча (28,32%). Примеры: 29 тур 2017/2018 Локомотив-Зенит 1:0 (рис.4); 2 тур 2018/2019 Зенит-Арсенал Тула 1:0 (рис.5) и 23 тур 2020/2021 ЦСКА-Зенит 2:3(рис.6). В первом случае xG и DEEP больше у победителя, но PPDA меньше, что говорит о разной тактике команд. У Локомотива качественные моменты (xG, 0,79 против 0,16) и они лучше проходили оборону (DEEP, 8 против 2), в то время как у Зенита была лучше интенсивность прессинга, но данная метрика не показывает успешность этого прессинга, и соперник легко проходил его, создавал моменты и забил 1 гол. Chances (60%) также подтверждает это. Во втором примере наблюдается равенство DEEP, при этом интенсивность прессинга Зенита почти в 3,7 раз выше, чем у Арсенала, а также у Арсенала практически не было голевых моментов - 0,14 xG, в то время как у Зенита этот показатель равняется 1,58. Это может говорить о том, что многие моменты были созданы со стандартов (свободный удар, угловой, пенальти и т.д.). Но если моментов действительно было мало (всего 2 DEEP, как еще, кроме как, со стандартов в данном случае создавать), то вес этих моментов был большой, т.е. выгодная позиция или ситуация. Арсенал испытывал трудности с продвижением мяча, исходя из PPDA, что не может не влиять на результат. В итоге вероятность победы в данном случае равна 84%. В последнем примере мы наблюдаем равенство голевых моментов у команд, но Зенит чаще входил в «штрафную» соперника, что сказывается на качестве единичных моментов, пусть даже их суммарная опасность равна. Лучше иметь немного, но качественных моментов. Однако итоговый Chances равен всего лишь 37% по сравнению с аналогичными примерами, несмотря на то, что статистика прямо пропорциональна результату. Это говорит о том, что метрика xG более показательна и совершенна.

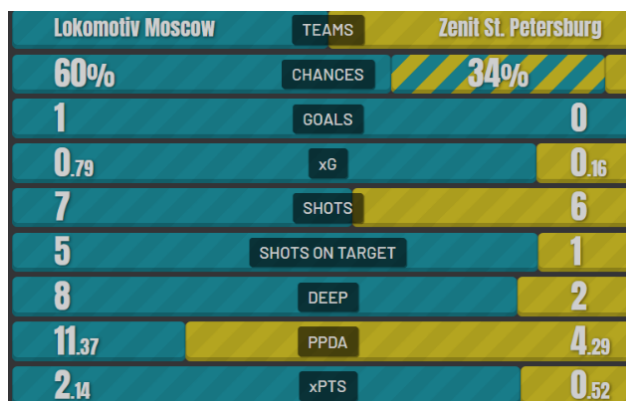


Рисунок 4 – Локомотив-Зенит

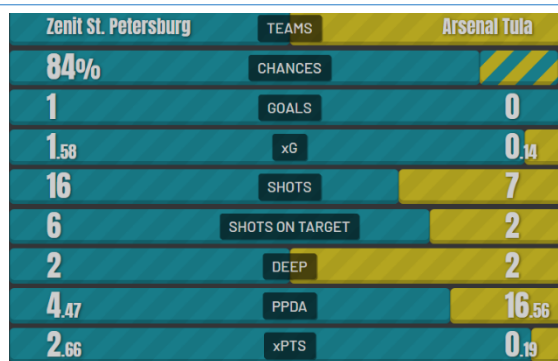


Рисунок 5 – Zenit-Арсенал



Рисунок 6 – ЦСКА-Зенит

Однако, не всегда аналитика позволяет правильно оценивать ситуацию на поле. Есть аномальные матчи, где статистика обратно пропорциональна результату - 23 матча (20,35%). Например, матч 21 тура сезона 2020/2021 Рубин – Зенит(табл.1). Зенит перебил соперника по всем показателям, но проиграл (Chances 5%). Со стороны кажется, что это невозможно, но таких матчей более чем достаточно. Рубин имеет очень низкий показатель интенсивности прессинга (13,71), в отличие от Зенита, у которого он очень высокий (2,89), что говорит о том, то Рубин намеренно отдал мяч Зениту, сел в низкий блок обороны и имел очень низкую среднюю линию обороны. Большие количество DEEP и вес xG не имеют смысла, так как многие из этих передач не были опасными в виду очень плотной обороны и большинство ударов были «мусорными». Рубин забил голы со стандарта и контратаки. Доминирование по ходу матча не всегда дает результат. Важнее грамотно выбранная тактика, тщательно проведенный анализ соперника, индивидуальные способности игроков и умелое использование своих шансов.

Таблица 1 – Рубин-Зенит(20/21)

xG	xG opp.	DEEP	DEEP opp.	PPDA	PPDA opp.	Итог	Chances
2,97	1,29	26	5	2,89	13,71	1:2	5%

Влияние xG. Есть случаи (8 матчей, 7,07%), когда характер матча является равным для обеих команд, но в конце матча устанавливается неравный счет, к примеру, как это было в матче 3 тура сезона 2018/2019 Рубин-Зенит(табл.2). Зенит превосходит соперника только по качеству моментов, а в остальном уступает ему, но этого достаточно для положительного исхода. В футболе не всегда решает красота или эстетика игры, иногда нужно сыграть прагматично и прямолинейно для достижения результата.

Таблица 2 – Рубин-Зенит(18/19)

xG	xG opp.	DEEP	DEEP opp.	PPDA	PPDA opp.	Итог	Chances
1,13	0,48	6	7	13,23	9,08	1:0	58%

Очевидно, что аналитика еще далеко не совершенна и её нужно модифицировать, делая корректировки, к примеру, на владение мячом или индивидуальные способности игроков, что уже развивается, особенно в ведущих европейских лигах. Аналитику также можно использовать для анализа индивидуальных особенностей игроков.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Kramsch C. The Political Power of the Algorithm // Technology and Language. 2020. № 1(1). С. 45–48. DOI:<https://doi.org/10.48417/technolang.2020.01.10>.
2. Быльева Д.С., Лобатюк В.В. Образ цифрового будущего общества через призму пандемии // Философская мысль. 2021. № 2. DOI: 10.25136/2409-8728.2021.2.35169
3. Бекетова Д.А., Савкина Н.В. "Информационные технологии в физической культуре и спорте" Наука-2020, № 2 (27), 2019, с. 78-81.
4. Expected Goals [Электронный ресурс]: определение, основы системы и методика расчета. URL: <https://www.sports.ru/betting/faq/1662919/> (дата обращения 04.04.2021).
5. Молдагереев М.М. "X-гид по футбольной аналитике" Вестник Московской международной академии, № 2, 2019, с. 54-63.
6. DEEP [Электронный ресурс]: визуализация метрики. URL: https://twitter.com/ru_stats (дата обращения 04.04.2021).
7. DEEP и PPDA [Электронный ресурс]: определение, методика расчета. URL: <https://www.sports.ru/tribuna/blogs/xgaltterpl/2099588.html> (дата обращения 04.04.2021).
8. Understat.com [Электронный ресурс]: URL: <https://understat.com> (дата обращения 04.04.2021).

Podkorytov Gleb,

Student,

Graduate School of Mechanical Engineering, Materials and Transport

Peter the Great St. Petersburg Polytechnic University

Saint-Petersburg, Russian Federation

FOOTBALL ANALYTICS

Abstract:

The article presents the results of the study of the correspondence of analytical data with the real results of soccer matches. Examples of the use of certain data, metrics to analyze the game and its result are considered.

Keywords:

Analytics, metrics, statistics, football.