

УДК 004.6:004.42

ПРИМЕНЕНИЕ ТЕХНОЛОГИИ BIG DATA В ИГРОВОЙ ИНДУСТРИИ



Д.В. Кишкевич
Студент БГУИР



С.Н. Нестеренков
Кандидат технических наук, доцент,
декан факультета компьютерных
систем и сетей



А. Н. Марков
Старший преподаватель,
магистр технических наук,
заместитель начальника
Центра информатизации и
инновационных разработок
БГУИР

Центр информатизации и инновационных разработок Белорусского государственного университета информатики и радиоэлектроники, Республика Беларусь
Белорусский государственный университет информатики и радиоэлектроники,
Республики Беларусь
E-mail: dkishkevich6@gmail.com, s.nesterenkov@bsuir.by, a.n.markov@bsuir.by

Д.В. Кишкевич

Студент 4 курса специальности “Вычислительные машины, системы и сети” БГУИР.

С.Н. Нестеренков

Кандидат технических наук, доцент, декан факультета компьютерных систем и сетей Белорусского государственного университета информатики и радиоэлектроники, доцента кафедры программного обеспечения информационных технологий. Автор публикаций на тему машинного обучения, алгоритмов принятия решений, искусственных нейронных сетей и автоматизации.

А.Н. Марков

Магистр технических наук, старший преподаватель кафедры ПИКС, заместитель начальника Центра информатизации и инновационных разработок Белорусского государственного университета информатики и радиоэлектроники.

Аннотация. Использование больших данных быстро превращается в общемировое явление. В данной работе анализируется использование больших данных в видеоиграх, технологии баз данных для решения задач игровой индустрии, и такие сервисы как Amazon Web Service. Нами рассматриваются примеры использования больших данных в видеоиграх. Проблемы разработки видеоигр касаются в основном игрового дизайна и количества различных характеристик.

Ключевые слова: Большие данные, miHoYo, Zynga, технологии NoSQL.

Введение.

Big Data или большие данные – это структурированные или неструктурированные массивы данных большого объема. Их обрабатывают при помощи специальных автоматизированных инструментов, чтобы использовать для статистики, анализа, прогнозов и принятия решений [1]. В современном мире Big Data – это современный тренд в аналитике, который позволяет компаниям принимать больше решений, основанных на данных, чем когда-либо прежде. Большие данные помогают разработчикам балансировать, тестировать и монетизировать свои игры. Крупные компании, такие как Microsoft, видят ценность агрегации данных и приобретают игровые компании, такие как Minecraft, за 2,5 миллиарда долларов, потому что они осознают

важность больших данных в долгосрочной перспективе и нуждаются в фирмах, передающих данные, чтобы помочь им добывать и понимать пользователей. По сравнению с популярными социальными сетями, которые обрабатывают большие объемы данных ежесекундно, игровым компаниям не обязательно использовать большие данные. Тем не менее, такие компании, как Electronic Arts и Riot Games, miHoYo используют большие данные для отслеживания игрового процесса, что помогает прогнозировать эффективность игры путем анализа 4 ТБ операционных журналов и 500 ГБ структурированных данных. Игровые данные поступают в нескольких формах. Electronic Arts, одно из крупнейших имен на рынке, ежемесячно проводит почти 50 миллиардов игровых минут в более чем 2,5 миллиарда игровых сессий [2]. Умножьте это на количество производителей игр и количество платформ, доступных сегодня. Огромное количество внутриигровых данных ожидает анализа. Теперь эти внутриигровые данные могут состоять из нескольких вещей:

1. Сбои в игре, которые были зарегистрированы в файлах журнала или о которых сообщалось. Разработчики игр могут воссоздать сценарий и работать над устранением таких проблем.

2. Игровые чаты и голосовые сообщения сохраняются и отслеживаются на наличие жалоб и проблем с игрой с помощью анализа настроений.

3. Взаимодействие между игроками в игре отслеживается с помощью файлов журналов, и разработчики могут со временем сделать их более реалистичными, анализируя тысячи или даже миллионы взаимодействий.

4. Данные о внутриигровых покупках или внутриигровые параметры, выбранные пользователями, или любые данные, сгенерированные пользователями в игре, то есть не во время самой игры, которые используются для понимания поведения конкретного пользователя и, следовательно, для продвижения целевой рекламы [3].

Далее будут подробнее рассмотрены некоторые способы использования больших данных при решении задач, с которыми сталкиваются разработчики и издатели.

Уравнивание шансов.

Балансировка одна из самых важных частей разработки игры, т.к. игра должна оставаться честной и захватывающей. Сохранить баланс между навыками разных персонажей остается довольно трудной задачей, из-за чего может создаваться определенное преимущество в игре одного персонажа над другими, так на пример нарушается принцип честной игры, а также это может влиять на выбор игроками определенного любимчика среди героев, что мешает получать задуманный разработчиками игровой опыт, иными словами игра может быстро наскучить. Для того, чтобы избежать подобного, во время игры собирается огромное количества статистики, например на основании информации о количестве побед и частоте выбора игроками определенного героя или оружия, разработчики могут принимать решение об усилении или ослаблении характеристик. Ярким примером такой игры, является Counter Strike: Global Offensive [4].

Однако это не единственный вид статистики, которую компания Valve собирает во время матчей. Для игр, основанных, на сетевой соревновательной системе актуальной является проблема сторонних приложений, использование которых, дает нечестное преимущество над другими игроками. Эту проблему также решают благодаря использованию больших данных. Записи матчей, количество сделанных убийств, скорость перемещения персонажа, количество найденных ресурсов, все эти данные бережно хранятся и в случае чего помогают определить нечестного игрока.

Балансировка игровой среды. Компания miHoYo выпустившая быстро набирающую популярность игру Genshin Impact нашла не совсем обычное применение большим данным в игровой индустрии. Каждый цикл обновления игрокам на игровую почту приходит письмо, где разработчики, за небольшое вознаграждение, предлагают игрокам пройти опрос. Опрос содержит вопросы о прошедшем цикле обновления [5]. Например, предлагается оценить

события прошедшего обновления или новый контент. Это позволяет очень оперативно реагировать на меняющиеся запросы игроков и в кратчайшие сроки изменять спорные моменты.

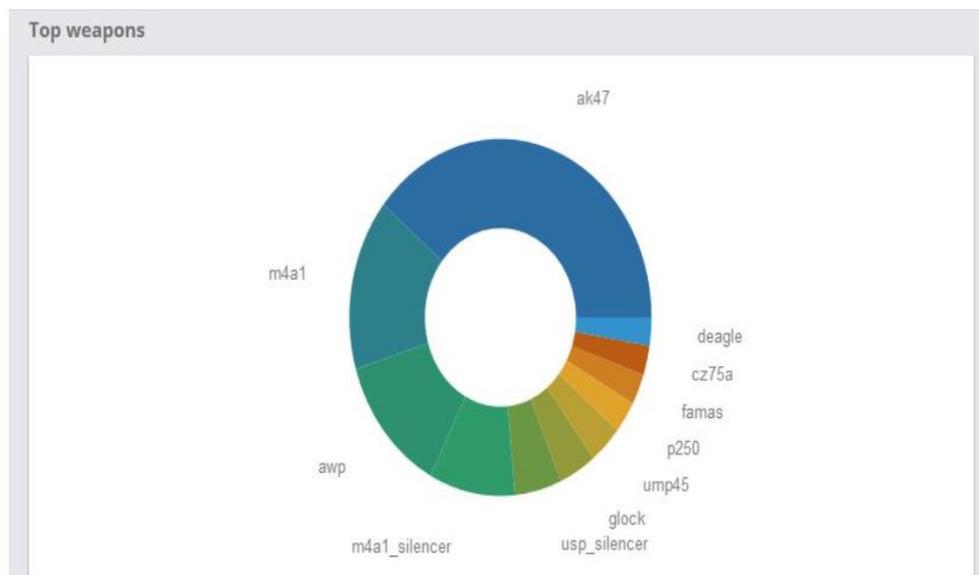


Рисунок 1 – Статистика с наиболее популярным оружием у игроков в игре Counter Strike: Global Offensive

На основании большого количества данных могут приниматься решения о продолжении игр или закрытии их. Существуют целые интернет-площадки, позволяющие пользователям оценить игру и высказать о ней свое мнение. Такие как Metacritic или Rotten Tomatoes. Эти площадки дают возможность игрокам влиять на игры посредством оценки, хоть и косвенно.

Технологии баз данных для решения задач в игровой отрасли. Одни из самых часто используемых баз данных являются реляционные базы, не смотря на начало их относительную старость, они все еще занимают весомое место в разработке игр. Реляционная база данных – база данных, основанная на реляционной модели данных [6]. Со временем появилась необходимость внедрения различных методов увеличения емкости таких баз данных.

Шардинг – метод разделения и хранения одного логического набора данных в нескольких базах данных. Разделение необходимо, если набор данных слишком велик для хранения в одной базе данных. Более того, многие стратегии сегментирования позволяют добавлять дополнительные машины. Разделение позволяет кластеру базы данных масштабироваться вместе с ростом его данных и трафика.

Денормализация – это процесс добавления предварительно вычисленных избыточных данных в нормализованную реляционную базу данных для повышения производительности чтения базы данных. Нормализация базы данных включает в себя устранение избыточности, чтобы существовала только одна копия каждой части информации. Денормализация базы данных требует, чтобы данные были сначала нормализованы.

Распределенное кэширование – это кэширование, которое используется совместно несколькими серверами приложений. Распределенный кэш может значительно повысить скорость отклика приложения, поскольку обычно данные извлекаются из кэша намного быстрее, чем из реляционной базы данных (или веб-службы).

Технология NoSQL позволяет обойтись без различных методов увеличения емкости. Базы данных NoSQL создаются с нуля для хранения и обработки огромных объемов данных в масштабе и поддержки растущего числа современных предприятий.

Преимуществами NoSQL являются:

1. NoSQL обеспечивает высокий уровень масштабируемости

2. Реализация менее затратна. Она обеспечивает хранение полуструктурированных данных, а также обеспечивает гибкость схемы.

3. К преимуществам NOSQL также относится возможность обработки:

- a. Большие объемы структурированных, полуструктурированных и неструктурированных данных.
- b. Объектно-ориентированные алгоритмы позволяют реализовывать их для достижения максимальной доступности в нескольких центрах обработки данных.
- c. Системы, основанные на согласованности событий, масштабируют рабочие нагрузки обновлений лучше, чем традиционные СУБД OLAP, а также масштабируются до очень больших наборов данных.
- d. Простое в использовании и гибкое программирование. Эффективная масштабируемая архитектура вместо дорогой монолитной архитектуры

Существует множество типов баз данных NoSQL: базы данных документов, хранилища вида «ключ-значение», базы данных, ориентированные на столбцы, и графические базы данных [7]. На данный момент применяется более 150 баз данных NoSQL, и их число продолжает увеличиваться. MongoDB имеет гибкую систему хранения, что означает: объекты для хранения не обязательно должны иметь одинаковые поля или структуру. MongoDB также имеет некоторые функции оптимизации, которые распределяют наборы данных, в целом являясь более сбалансированной и ориентированной на производительность системой. Другие системы баз данных NoSQL, такие как Apache CouchDB, также относятся к типу хранилища документов и имеют много одинаковых функций с MongoDB, с добавлением того, что для получения доступа к базе данных можно использовать RESTful API. REST – это архитектурный стиль для обеспечения стандартов между компьютерными системами в Интернете, упрощающий взаимодействие систем друг с другом. Системы, совместимые с REST, часто называемые системами RESTful, характеризуются тем, что они не имеют состояния и разделяют задачи клиента и сервера. Одной из самых популярных баз данных NoSQL является Cassandra, разработанная Facebook. Cassandra – это база данных хранилища столбцов, которая включает в себя множество функций, направленных на надежность и отказоустойчивость.

Монетизация. В эпоху организаций, управляемых данными, каждая игровая компания пытается собрать все доступные данные в своих хранилищах данных. В большинстве случаев собираемые данные касаются игроков и их деятельности. Из приложения для мобильных игр можно было собирать данные обо всех игровых действиях клиентов. Собранные точки данных – прогресс, лайки, взаимодействия, предпочтения, расходы, демография, модели поведения и многие другие данных. Кроме того, такого рода информация часто используется для получения данных в игровой индустрии .

Среди огромного количества транзакционных данных выделяются наиболее важные факторы, зависимости, корреляции и создаются интеллектуальные алгоритмы для получения наибольшей прибыли. Ведь стать прибыльной компанией без хорошей модели монетизации нельзя. Большие данные могут быть очень полезны для модели монетизации F2P, потому что экономическая архитектура A-R-M-D, дает возможность улучшать монетизацию игры постоянно, используя данные игрока в процессе игры. Zynga стала одной из первых компаний, использующих разработку игр при помощи данных в своих странах. Член-учредитель Андрей Трейдер сказал в интервью The Wharton School: «грязный маленький секрет Zynga – из пяти корпоративных ценностей ни одна не является более важной, чем управление метриками. Для Zynga это означало, если вы не можете измерить что-то, не стройте его». Zynga внимательно следила за охватом, удержанием позиций и доходом при этом удержание является наиболее важным фактором. Их доход был получен за счет продажи внутриигровых предметов, которые хотели их клиенты [8].

Заключение.

Сегодня, большие данные широко используются в игровой индустрии, использование технологий сбора данных значительно упрощает работу разработчикам, а также улучшает

игровой опыт пользователей. Значительную долю современных баз данных, занимают довольно старый Реляционные базы данных, однако современные системы стремительно приходят на рынок. Большие данные принесли больше в таблицу разработчиков игр с точки зрения балансировки и монетизации. Многие решения в игровой индустрии принимаются, опираясь на сырую статистику, полученную в результате сбора данных.

Список литературы

- [1] Bernard, M. Big Data: Using SMART Big Data, Analytics and Metrics To Make Better Decisions and Improve Performance / M. Bernard. – London : John Wiley & Sons Ltd, 2015. – 259 с.
- [2] Вольф Марк, Дж. П. Энциклопедия видеоигр: Культура, технология и искусство игры. / Дж. П. Вольф Марк. – Greenwood Press, 2012. – 763 с.
- [3] Базы данных AVS для игр. [Online]. Available: <https://aws.amazon.com/ru/gametech/databases/> [Accessed: 27 March 2022].
- [4] HLTV.ORG . [Online]. <https://www.hltv.org/> [Accessed: 27 March 2022].
- [5] About game – Genshin Impact. [Online]. Available: <https://genshin.hoyoverse.com/ru/game> [Accessed: 27 March 2022].
- [6] Кузин, А.В. Базы данных. Учебное пособие для студентов высших учебных заведений. / А. В. Кузин. – Второе издание, изд. центр «Академия», 2008. – 320 с.
- [7] Фаулер, М. NoSQL: новая методология разработки нереляционных баз данных / М. Фаулер, Дж. Садаладж Прамодкумар. – Москва : Вильямс, 2013. – 172 с.
- [8] K. Watanabe, T.Fukamachi, N.Ubayashi and Y.Kamei, Poster: Automated A/B Testing with Declarative Variability Expressions, IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW), IEEE 2017.
- [9] Нестеренков, С.Н. Адаптивный поиск вариантов расписания с использованием модифицированного генетического алгоритма / С.Н. Нестеренков // Вести Института современных знаний. - 2015. - N 2. - С. 67-74.
- [10] Нестеренков, С.Н. Функциональная модель процедур планирования и управления образовательным процессом как основа построения информационной среды учреждения высшего образования / С.Н. Нестеренков, Н.В. Лапицкая // Вести Института современных знаний. - 2018. - N 1. - С. 97-105.
- [11] Нестеренков, С.Н. Сетевая модель и алгоритм составления расписания учебных занятий на основе данных прошлых периодов / С.Н. Нестеренков, Н.В. Лапицкая, О.О. Шатилова // Вести Института современных знаний. - 2018. - № 4. - С. 85-92.

APPLICATION OF BIG DATA TECHNOLOGY IN THE GAME INDUSTRY

D. V. KISHKEVICH

*Student of BSUIR,
Software engineer*

S.N. NESTERENKOV

*PhD, Associate Professor, Dean of the
Faculty of Computer Systems and
Networks*

A. N. MARKOV

*Senior lecturer of the
department, Deputy head of the
Center for Informatization and
Innovative Developments*

Center for Informatization and Innovative Developments of the Belarusian State University of Informatics and Radioelectronics, Republic of Belarus

Belarusian State University of Informatics and Radioelectronics,

The Republic of Belarus

E-mail: dkishkevich6@gmail.com, s.nesterenkov@bsuir.by, a.n.markov@bsuir.by

Annotation. The use of big data is rapidly becoming a global phenomenon. This paper analyzes the use of big data in video games, database technologies for solving the problems of the gaming industry, and services such as the Amazon Web Service. We consider examples of the use of big data in video games. The problems of video game development are mainly related to game design and the number of different features.

Keywords: Big data, miHoYo, Zynga, NoSQL technologies.