

## Введение

«BitReconstructor» (уровень «BR-L1») представляет собой программный комплекс, предназначенный для восстановления файла из набора «недостоверных» копий, путем применения «бинарного голосования». Предполагается, что «BitReconstructor» не заменяет существующие программные комплексы для восстановления данных, а лишь дополняет их, реализуя простейшую статистическую коррекцию на уровне отдельных бит. «Бинарное голосование» - это процесс выбора одной из гипотез значения бита «0» или «1», основанный на сравнении накопленных вероятностей (таким образом, выбирается наиболее вероятное состояние). Вероятности накапливаются в соответствии с весовыми вкладами источников (т.к. оценить истинный вес, т.е. доверие к источнику достаточно сложно, все источники имеют одинаковый вес, одинаковую значимость вносимых в голосование данных).

Программный комплекс состоит из двух приложений: консольного «BitReconstructor.BinVote.Console.exe» и GUI-версии «BitReconstructor.BinVote.GUI.exe». Общая схема применения программного комплекса очень проста: на вход программы подаются копии одного и того же файла (с разными именами), а на выходе формируется «более достоверная» копия (на основе входных файлов, имеющих наибольший из наиболее часто встречающихся объемов).

GUI-версия позволяет обрабатывать входные наборы до 5 файлов, и если количество «кандидатов» больше, рекомендуется использовать консольное приложение. В GUI-версии входные файлы-кандидаты автоматически проверяются на доступность (проверяется гипотеза «файл существует»). Кроме того, для указанного набора кандидатов производится предварительная фильтрация, через те же самые механизмы и идентично тому, как это будет сделано при запуске обработки (в обработку пропускаются файлы, имеющие наиболее часто встречающийся наибольший одинаковый объем).

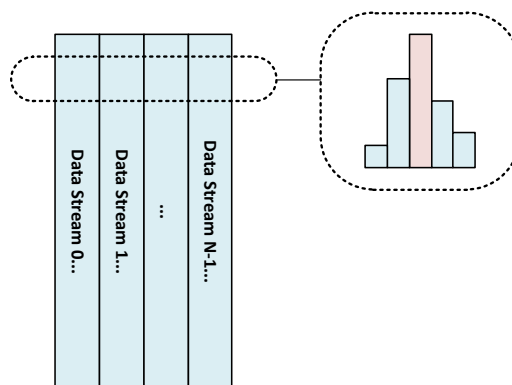


Рисунок 1 – Бинарное голосование «BR-L1»

## Применимость утилиты «BitReconstructor»

Утилита «BitReconstructor» абсолютно точно может помочь Вам, если Вы имеете 3 или более копий поврежденных данных. Такими копиями являются файлы, в том числе и iso-образы (или любые другие bin-файлы, содержащие идентичную битовую структуру). То есть, например, имея набор поврежденных аудиодисков, можно снять iso-образы с трех или более поцарапанных болванок (при помощи одной и той же программы, у которой активирован пропуск «битых» секторов) и «прогнать» имеющиеся потоки через указанную утилиту. Или же, имея набор поврежденных архивов на «флешках», можно скопировать их в одну папку под разными именами, либо, используя GUI-версию, указать пути к ним прямо на исходных носителях, и запустить процесс восстановления данных. На вход программы можно подавать любые файлы, в том числе exe, msi, doc, docx, xls, xlsx, jpg, bmp... Самое главное – чтобы входные кандидаты являлись копиями одной и той же битовой структуры.

Чем больше потоков на входе, тем лучше. Однако, имеется разумный предел. По-моему мнению, оптимальное количество потоков – 5, а базовое – 3. Впрочем, никто не запретит Вам сформировать 9, 11 или более iso-образов, и, применяя консольное приложение, получить выходной файл с ещё большей достоверностью (чем больше кандидатов участвует в бинарном голосовании, тем лучше).

Древовидный принцип наложения «патчей» также применим. Имея 9 копий, можно осуществить 3 прогона по 3 файла, с объединением в заключительном «прогоне» результатов предыдущих запусков (но, например, имея 11 файлов, осуществлять 2 прогона по 5 файлов, с подключением на заключительном этапе 1-го не входившего в обработку – неверно, т.к. он, имея меньший «класс точности», будет являться своего рода «помехой»). Иными словами, в случае многоэтапного голосования, на вход утилиты нужно подавать лишь равных по достоверности кандидатов, т.к. имеющих равное количество предков в голосованиях. А ещё лучше – не применять многоэтапных схем. При наличии большого количества копий – сразу же переходить к консольной версии.

Если у Вас имеются поврежденные RAR-архивы с данными для восстановления и их корректировка на основе данных для восстановления невозможна (слишком много битовых ошибок), «BitReconstructor» может дать им шанс. Если Вы имеете достаточное количество копий каждого архива (3 или более), существует вероятность, что «BitReconstructor» сможет исправить большую часть битовых ошибок (или вообще откорректировать данные на все 100%). Оставшиеся ошибки, если таковые будут, могут быть исправлены как раз при помощи упомянутых данных для восстановления внутри результирующего архива (полученного на основе бинарного голосования).

Утилита «BitReconstructor» может также использоваться для повышения достоверности двоичных данных. Предположим, ставится задача скопировать содержимое какого-либо носителя, имеющегося в трех или более копиях, так, чтобы полученная копия или их набор, был максимально достоверным. Следует снять образы с исходных носителей (iso, bin и т.д.) и «прогнать» их через «BitReconstructor». Полученный файл-образ (или просто файл) будет обладать максимальной достоверностью среди своих предков, и именно с него и нужно формировать новые копии.

## Тест на простом примере (часть I)

Для того, чтобы показать, как работает приложение, создадим «подопытного кролика». Сформируем 5 копий этого файла и внесем в каждую копию повреждения так, чтобы в пересекающихся позициях было не более 2 ошибок из 5 возможных. Данный тест прилагается к распространяемому пакету ПО и документов.

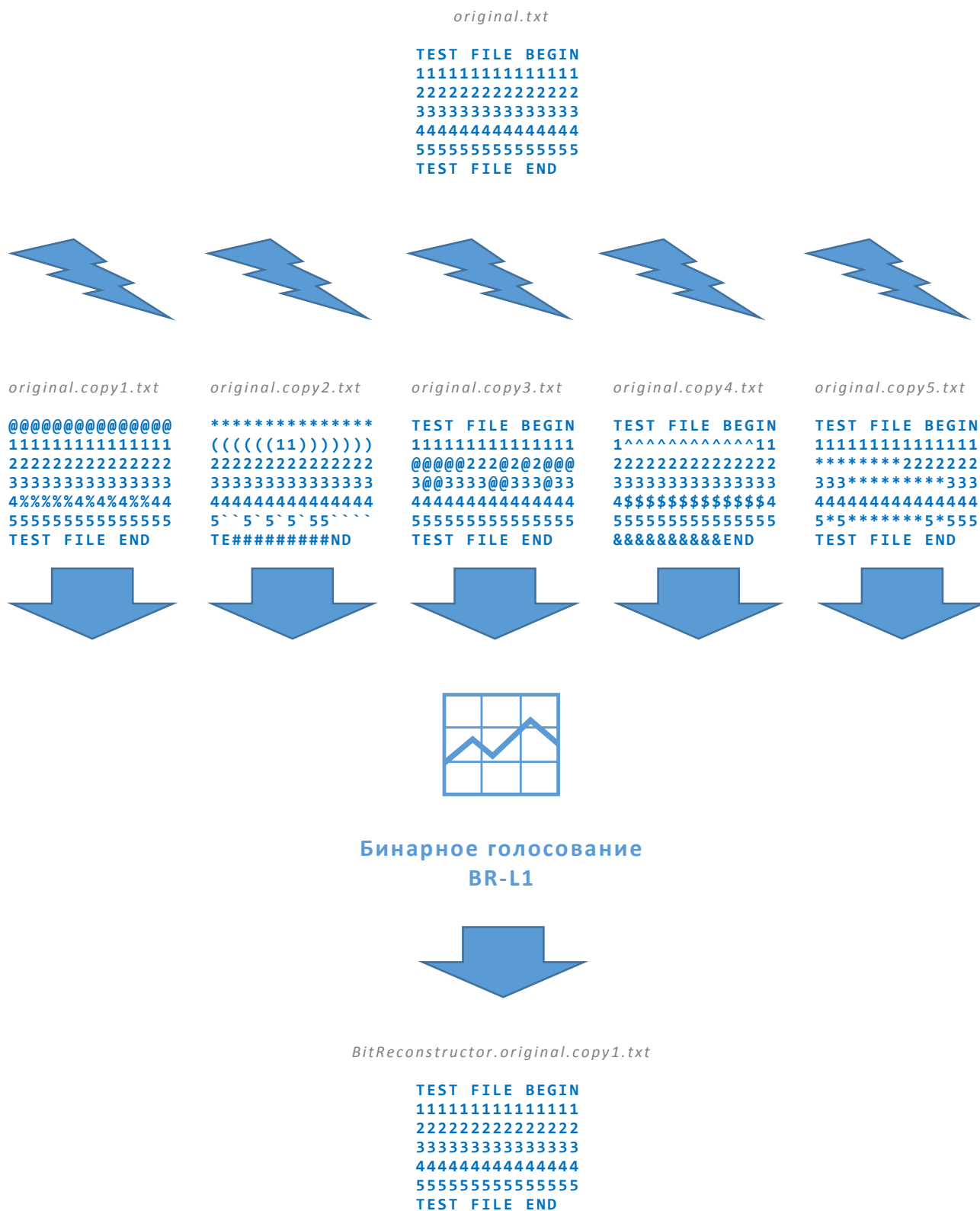


Рисунок 2 – Пример работы схемы бинарного голосования (5 потоков)

## Тест на простом примере (часть II)

Запустите приложение «BitReconstructor.BinVote.GUI.exe» и последовательно укажите путь к каждому из файлов-кандидатов, активируя выбор файла нажатием на экранную кнопку >>. После выбора каждого из файлов, фон текстового поля указывает на результат предстартового теста. Если поле имеет красный фон, файл не существует, желтый фон указывает на доступность файла, но пока не подтвержденную алгоритмом проверки возможность включения в обработку, зеленый фон сообщает о возможности корректного включения указанного файла в процесс.

После того, как предстартовый тест набрал минимально-возможное количество файлов, становится активной экранная кнопка «Go / Старт». Если не указать имя выходного файла, оно будет сформировано автоматически, по имени первого входного потока, к которому будет добавлен префикс «BitReconstructor.».

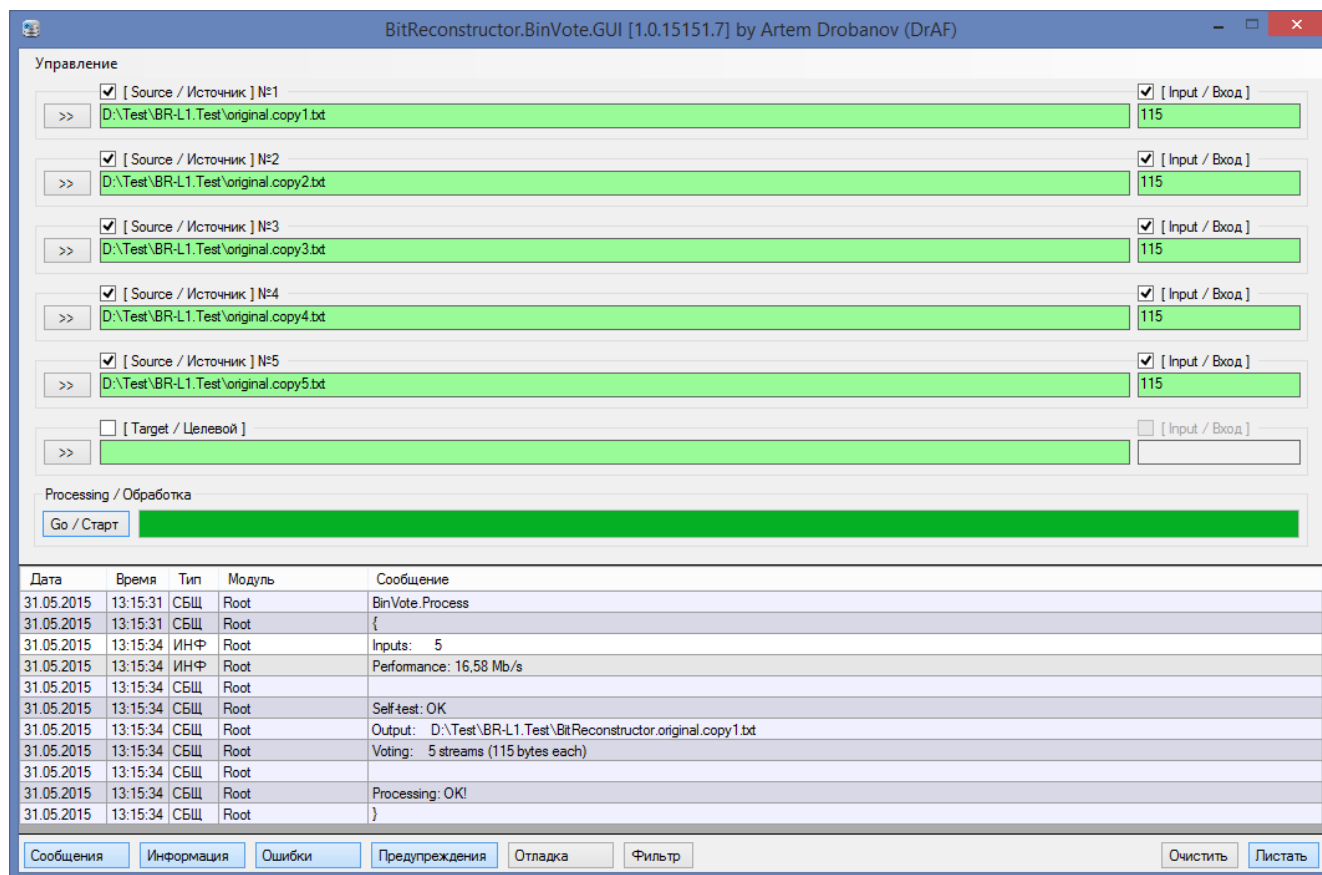


Рисунок 3 – Пример запуска схемы бинарного голосования в GUI-приложении

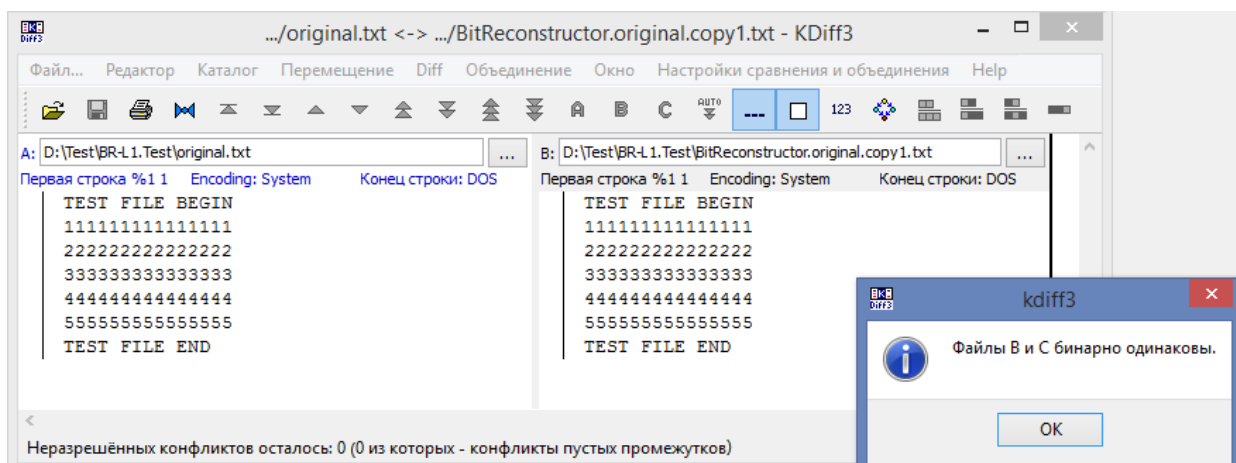


Рисунок 4 – Проверка результатов бинарного голосования (GUI)

### Тест на простом примере (часть III)

Осуществим теперь аналогичный процесс через консольное приложение. Запустим файловый менеджер «Far», скопируем исполняемые файлы консольного приложения в рабочую директорию и сформируем командную строку:

```
«BitReconstructor.BinVote.Console.exe original.copy1.txt original.copy2.txt  
original.copy3.txt original.copy4.txt original.copy5.txt OUT.txt»
```



Рисунок 5 – Пример запуска схемы бинарного голосования в консольном приложении

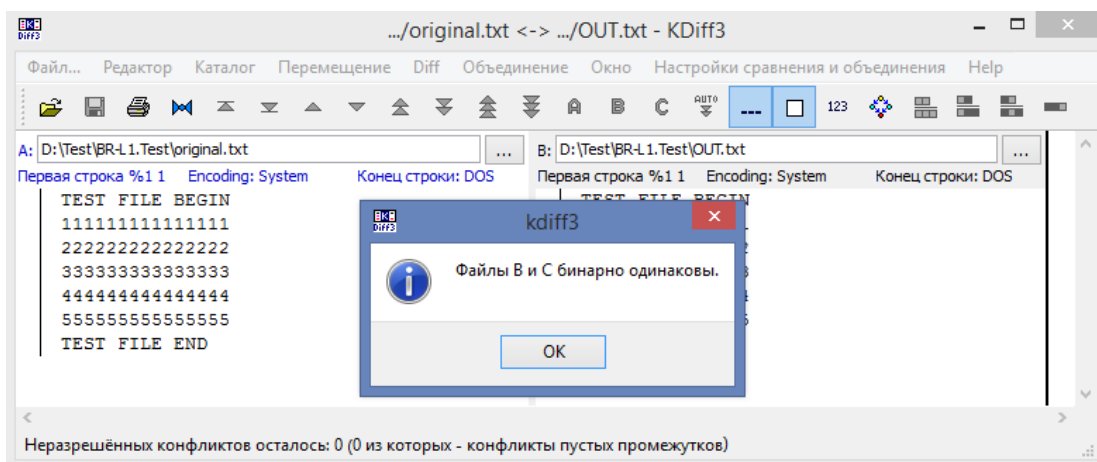


Рисунок 6 – Проверка результатов бинарного голосования (консольное приложение)