Title: PROCESSING OF BROADCASTING SCHEDULES

Abstract: The invention provides a method of processing data to enrich a broadcast schedule (16) of titles, times and channels of scheduled broadcasts of television programmes. The method includes allocating a program identity to each program, identifying programmes that form episodes of series and non-series programmes and allocating an episode identity to each episode in a series. Episodes that are identified as repeats are referenced chronologically in relation to the current premiere episode according to the offset value and are assigned the relevant existing episode identity. The data is then saved in an enriched program schedule (18) that can be transmitted to subscribers (14) to allow them to program PVRs to record the programmes in the schedule (18).
PROCESSING OF BROADCASTING SCHEDULES

FIELD OF THE INVENTION
This invention relates to the processing of program schedules to allow selection of programmes for recording on a personal video recorder (PVR) or any like recorder for programmes broadcast via conventional television, the internet, mobile telephone networks, or the like.

BACKGROUND TO THE INVENTION
PVRs, also known as digital video recorders (DVRs), are devices that record video to hard drive-based digital storage media, which may require encoding analogue signals to store them in digital formats, or directly storing them in the cases of digital broadcasts. Many subscription television broadcasters, such as cable or satellite television broadcasters, incorporate PVR functions in their set-up boxes.

Recording digital incoming subscription television broadcasts in a PVR allows subscribers to access and manipulate this data in a number of ways, including pausing live television, instant replays, skipping advertising, or the like. Of particular importance to the present invention, is the ability of a subscriber to program a PVR to record selected broadcasts, especially to record a series of programmes or a program that forms part of a series (i.e. an episode).

A broadcaster of digital satellite television typically receives the contents of programmes that it will broadcast, from outside suppliers who identify the programmes in a variety of ways, e.g. by providing programme titles, series titles, episode titles, series or episode numbers, or the like. These designations are not used consistently between different suppliers of the program contents and/or are not used consistently by individual suppliers. The identification of programmes is further complicated by the fact that some programmes that form part of a series, are not identified as such and when programmes form a series, the number of episodes that make up the series, is not known. It is also common for different series of a program to be broadcast concurrently on different channels and/or in different time slots. In
some instances, such as the broadcasting of the news or live sports events, the contents and duration of programmes is often not known in advance.

The broadcaster arranges the programme contents in advance of broadcasting and makes a schedule of planned broadcasting available to its subscribers, typically in print, but in some cases electronically. A subscriber with a PVR now has the ability to program the PVR to record some programmes selectively. In the simplest form, the PVR can be programmed to record the broadcast on a predetermined channel at predetermined times, according to the schedule of the broadcaster. Many PVRs also have the capability to identify the selected programmes and thus not to start recording at the exact time for which it has been set, but to time the recording to the start and finish of a program.

The programming of a PVR in this manner has a number of deficiencies in that it does not allow the flexibility of recording the most convenient one of different (repeat) screenings of episodes, to avoid a clashing broadcast on another channel. It does not assist the subscriber in identifying the program contents, to avoid duplication of recordings, avoid missing some broadcasts, etc. The programming can be tedious, e.g. when a number of episodes of a series appear at different time slots, on different days, etc.

These disadvantages could be alleviated if a broadcaster can present the program transmission schedule to the subscriber, in such a way that it is easy for a subscriber to select programmes, especially if the subscriber can identify programmes to be recorded by simply identifying a series of which the programmes form part. However, volume and complexity of program transmission schedules can make the task of preparing such a schedule extremely difficult and time consuming.

The present invention seeks to provide a method for enrichment of program transmission schedules to allow convenient programming of a PVR to record selected programmes, especially programmes that form part of a series.
SUMMARY OF THE INVENTION

According to the present invention there is provided a method of processing data, said data comprising titles, times and channels of scheduled broadcasts of television programmes, said method comprising:

- allocating a program identity to at least some of the programmes;
- identifying programmes that form episodes of series and non-series status of programmes that do not form episodes of series;
- allocating episode identities to at least some of such identified episodes; and
- saving the data in a program schedule, said schedule including the program identities of at least some programmes, the series and the episode identities of at least some of the programmes that form episodes of series and the non-series status of at least some programmes that do not form episodes of series.

The program identity of at least one program may be determined by searching for a corresponding program title within the data being processed and on a historical database and/or by removing characters from the program title and then searching for a corresponding program title within the data being processed and on a historical database.

Programmes that form episodes of series are identified by grouping programmes with similar allocated identities, together and/or by searching within the data being processed and on a historical database, for a series with a title corresponding to a program title.

The method may include determining the frequency with which the episodes of a series are scheduled to be broadcast, by determining the order and times when episodes are scheduled to be broadcast, and assessing from the scheduled broadcast times and order, in accordance with predetermined criteria, whether a scheduled broadcast will be a first broadcast or a repeat of an episode.

The predetermined criteria may include a ranking of broadcast times.

The method may include reviewing the scheduled broadcast times of
programmes with similar identities and identifying the programmes as forming part of a series if the scheduled broadcasts of at least some of the programmes are in close succession in a back-to-back configuration.

The number of programmes may be determined, which have similar program identities and which are scheduled to be broadcast within a predetermined time slot, and the programmes may be identified as forming episodes of a series if the number of such programmes exceeds a predetermined number.

A series of programmes that is scheduled to be broadcast after a predetermined offset period during which no other program was broadcast with the same program identity as the programmes of the series, may be identified as being the first broadcast of the series, i.e. the premiere.

The method may include episode identities provided in the unprocessed data, of episodes scheduled to be broadcast before a predetermined cut-off date, to identify the episodes.

The method may include determining whether an episode is being repeated, by searching within the data being processed and on a historical database for episode titles corresponding to the episode identities.

Episodes that are being repeated may be identified by checking if they are scheduled to be broadcast within a predetermined period of time (i.e. the frequency) after the premiere episode. Episodes that are being repeated may also be automatically determined based on the timeslots within which they are scheduled to be broadcast.

A repeat offset may be determined automatically by comparing the broadcast times of the premiere of the series with the broadcast times of the first repeat of the series and the repeat offset may be used to identify episodes of a series chronologically in relation to the premiere of the series.
The method may be automatically performed on a computer or other such electronic device, but data on the schedule may be manually manipulated, e.g. to prevent the automatic processing of the data.

The invention extends to a set of instructions, stored on data storage means and being configured to be executable on a computer or other such electronic device, to perform a method of processing data as described hereinabove.

The invention further extends to a system for processing data, said system comprising data storage means on which a set of instructions is stored for performing a method of processing data as described hereinabove, and a computer or other such electronic device that is configured to perform the method of data processing, when the set of instructions is executed on the computer or other such electronic device.

BRIEF DESCRIPTION OF THE DRAWINGS

For a better understanding of the present invention, and to show how the same may be carried into effect, reference will now be made, by way of non-limiting example, to the accompanying drawings in which:

Figure 1 is a diagrammatic representation of the treatment of data relating to programmes for broadcasting, in accordance with the present invention;

Figure 2 is a flow diagram of the overall enrichment process in accordance with the present invention;

Figure 3 is a flow diagram of the "Find/Create Program ID" part of the enrichment process;

Figure 4 is a flow diagram of the "Determine Series Status and Frequency" part of the enrichment process; and

Figure 5 is a flow diagram of the "Find/Crate Episode ID" part of the enrichment process.
DETAILED DESCRIPTION OF THE DRAWINGS

Referring to Figure 1, a broadcaster 10 receives programmes from suppliers 12 for broadcasting to its subscribers 14. For the sake of simplicity, only one supplier 12 and one subscriber 14 are shown, to represent much larger, indeterminate numbers of these persons.

The broadcaster 10 arranges the programmes it receives from the suppliers 12 to form a schedule 16 and uses the present invention to enrich the schedule to obtain an enriched schedule or guide 18, that is made available to the subscribers 14, e.g. by transmitting it to them. The enriched schedule 18 is stored on a database 20 and data on the database relating to previous enriched schedules, is accessed during the enrichment process.

The data contained in the schedule 16 typically includes the names or titles that the suppliers 12 have allocated to programmes, as well as the intended times and channels for broadcasting the programmes and other bibliographic details such as episode names, category identifiers, or the like, but these details may be incorporated in the program titles or may be omitted. For reasons provided above, the data in such a schedule 16 has limited use for a subscriber in programming a PVR to record selected programmes, especially if the subscriber intends recording programmes that form part of a series and/or if the programmes intended to be recorded are interspersed with other programmes on other channels, which are to be recorded. The purpose of the enrichment of the schedule 16 is to identify in the enriched schedule 18, each program, and any series of which the programmes may form part, as well as the episode in such a series.

The enrichment method of the present invention allows for the data of the schedule 16 to be edited in a systematic manner, so that it can be automated. An overview of the enrichment process is shown in Figure 2, which includes three main modules that are completed in sequence and starts with the uploading of settings for the process. The first module 22 loops through each program that is
scheduled to be broadcast on each day, to find or create an identity for each program in the schedule 16 by applying a method that is described below in more detail, with reference to Figure 3. The second module 24 loops through all the programmes scheduled to be broadcast on each channel, grouping the programmes according to their identities to identify series of programmes and determine the frequency by which episodes in a series are scheduled to be broadcast, by applying a method that is described below in more detail, with reference to Figure 4. The third module 26 loops through each episode of a series scheduled to be broadcast on a day and loops through each day for which broadcasting has been scheduled to find or create an identification for each episode in the schedule 16, by applying a method that is described below in more detail, with reference to Figure 5.

Once the methods of each of the modules 22, 24 and 26 have been completed, the program identities, the series statuses and frequencies and the episode identities are included in the schedule and/or replace the corresponding information in the schedule, to provide the enriched schedule 18. This final enrichment process is identified in Figure 2 by the method step: "process guides".

Referring to Figure 3, in the first module 22, the name allocated to each program by its supplier 12 is determined. These names are often inconsistent with names previously used to identify the same program or series of which it forms part, which data would be stored on the database 20. The database is thus first searched to find a program name corresponding to the name allocated by the supplier and if a match is found, the matching name is presumed to be the correct name for the program and is assigned to the program as its identification name, which implicitly gives it a reference to the earlier program by the same name. Relevant bibliographic details of the program are retained on the database.

If the name allocated to the program by the supplier 12 does not match any entry on the database 20. Characters and/or words are stripped or removed from the allocated program name before comparing the name to the database for a second time. The removal of characters takes place according to preferences,
configured to delete obsolete details often found in names allocated by suppliers. If the stripped name still does not match any program name on the database, the stripped name is assigned to the program and the stripped and original names are saved on the database as its identification. If the stripped name matches a name in the database 20, the matching (striped) name is assigned to the program as its identification and is saved on the database, together with its original name. The saving of the original names allows for easier identification of programmes in future, which may have the same original names, thus avoiding having to strip the program names first, to identify similarities with previously identified programmes.

Referring to Figure 4, in the second module 24, to determine series status and frequency of episodes in a series, the programmes are grouped together according to their identifications (i.e. grouped together if they have the same names as determined in the first module 22). Known program settings are loaded from the database 20. These settings are in some instances configured to allow the settings to be overridden or to lock them, i.e. to prevent them from being overridden. If the settings for a program are locked, the particulars of that program as they appear on the database are presumed to be correct and no further action is taken in respect of the particular program in the second module.

If the settings of a program are not locked, its identity is compared to the data from the database 20, to determine whether or not it corresponds to the identity of programmes in a known, previously identified series. If the program identity corresponds to a known series, the settings of the known series are checked to see if programmes in the known series have been overridden as a non-series (e.g. when programmes have been found to have the same title, but to be unrelated). If the programmes with the corresponding identity have previously been overridden as a non-series, this is presumed to be correct in respect of the new program with the same identity and it is saved on the database 20 as a non-series program. Otherwise, if the known series of programmes with the same identity have not been overridden as a non-series, the new program under assessment is presumed to form part of that series.
If programmes with the same identity do not form part of a known series (i.e. they appear to be an unknown series), their other bibliographic data are checked. If they have category listings that list them as a series or mini-series, they are presumed to be a series. If they do not have such category listings, their episode names are checked to see if they have different episode names, to determine whether or not they form a series.

Once all the programmes with the same identities have been assessed to determine whether or not they are locked and to form a presumption of whether or not they are part of series, the module 24 loops through all the channels for each specified time slot in the broadcast schedule, searching for any episodes, i.e. any programmes of a presumed series, which may be scheduled to appear in the time slot. The time slots may be of any convenient duration, e.g. two hour time slots and are looped in a predetermined order that is specific to the channel. For most channels, a time slot of 19h00 to 21h00 may be the slot during which the first broadcast of a program is most likely to occur and this slot would be searched first, followed by other time slots in a decreasing likelihood of being the time slot of first broadcast (première). If an episode is found, the first day of the week and the earliest time during the particular time slot when the program is planned to be broadcast, is identified. If more programmes with the same identity are planned to be aired in a time slot, the scheduled broadcast times for the programmes are compared to see if they are intended to be broadcast in close succession (i.e. "back-to-back"), in which case they are identified as a series on the database.

The number of episodes of each series that are planned in the schedule to be broadcast (aired) within a time slot is determined, with some degree of allowance or padding. The allowance is affected by counting the number of episodes scheduled to be broadcast during a period starting at a predetermined time before the earliest starting episode during the time slot and ending at a predetermined time after the latest ending episode during the same time slot. If the interval between episodes (e.g. daily, weekly) within the time slot is greater than the
predetermined number, episodes with a regular interval are considered to form a series. If the programmes are found to form a series, the first planned broadcast in a given time slot (such as a week) of an episode of the series, is recorded, as well as the frequency (i.e. interval between episodes) with which episodes of the series are scheduled to be broadcast and the time slots and repeat offsets of scheduled repetitions of episodes are determined according to predetermined criteria and are stored. The repeat offset chronologically defines episodes in relation to the current premiere episode being aired, the offset can be automatically determined by comparing the first repeat of a new season to the first aired premiere of the new season.

Referring to Figure 5, in the third module 26, a unique identity is allocated to each program, to make it possible to identify each non-series program as well as each episode that forms part of a series, before updating the database 20 with the information attributed to each program or episode to provide the enriched schedule 18.

If a program was found in the second module 24, not to form part of a series, its details on the database 20 are updated, to be included in the schedule 18.

The historical data on the database 20, relating to the series of which the episode under consideration forms part, is checked and if no matches for this episode are found, it is presumed to be a new episode, different from any episode previously broadcast and is saved on the database as a new episode, in the form of a new episode entry. The comparison between the existing data on the database 20 and new programmes, is intended to avoid re-assigning of program or episode data, to programmes or episodes for which data already exist.

If the episode under consideration has scheduling that matches an earlier entry on the database 20, exactly, e.g. if it is planned to be broadcast at the same time on the same day of the week, the episode is presumed to be the same as the earlier episode and schedule 18 is updated with the known identity. Similarly, if
the episode under consideration has scheduling that matches an earlier entry, given padding or a tolerance of a predetermined time period difference between the times of the different episodes, the episode is presumed to be the same as the earlier episode and schedule 18 is updated with the known identity.

In this regard it should be borne in mind that the method of the invention can be performed repeatedly, each time relating to schedules of future broadcasts that may overlap, with the result that broadcasting schedules including the same programmes and episodes may be under consideration a number of times, before the broadcast dates. A typical practical application of the invention would include repeating the method daily or weekly, each time in relation to a schedule of programmes for the full month ahead (or for however far the schedule extends into the future). This allows changes to the schedule 16 that may be effected from time to time, to be incorporated in revised schedules 18 that are sent to subscribers from time to time.

A channel may have a specified cut-off date indicating the validity of the episode identities in schedule 16 as provided by the supplier 12 within a given period of time. If the episode falls within the cut-off period it, its supplied episode identity is considered to be valid and is used as the episode name when cross checking against the historical database. If the episode occurs after the cut-off period, the supplied episode identity is discarded. The episode’s name is checked to see if it matches an existing episode name and if so, schedule 18 is updated with the known episode identity.

If the episode is scheduled in close succession with other episodes of the same series, i.e. if the episodes of the series are scheduled to be broadcast back-to-back, this characteristic is saved on the database 20 and the episode is considered to be a unique and have its own identity.

The rule is applied that the first program to be scheduled in the predetermined, most likely time for a premiere, is presumed to be the premiere of an
episode and following episodes with the same episode name or within a predetermined period of time (i.e. the frequency) or in a specified time slot, are presumed to be repeats of the episode. This is referred to in Figure 5 as the "first time and first day" rule. If a scheduled transmission of an episode is found by application of this rule, to be a repeat of an earlier transmission of the same episode within the offset period, this information is saved to schedule 18.

Once the database 20 has been updated with all the latest data, including revisions, resumptions, references and the like, the database is checked to see if the planned broadcast of the episode will be the "first run", i.e. the first broadcast of the episode on the particular channel, or whether it has been previously shown. The database 20 is then updated with all the relevant data generated during the enrichment process, to provide the schedule 18.

Those skilled in the art will appreciate that all the steps in the method as illustrated in Figures 2 to 5 can be automated to be performed on a computer, allowing the method to be performed on a very complicated schedule with many repetitions of programmes or episodes, etc. The method is typically stored on electronic storage means such as a computer's memory or a data storage device, as software comprising a set of instructions that can be executed on the computer, to perform the method.

Once the subscriber 14 has received the enriched schedule 18, he can select to record all the episodes in a series of programmes as and when they will be broadcast and the data in the schedule 18 will allow fully automated recording of the episodes, since the episodes, the times and channels of their planned broadcast and the scheduled broadcasts of repeats of the episodes, are all determinable form the schedule.
CLAIMS

1. A method of processing data, said data comprising titles, times and channels of scheduled broadcasts of television programmes, said method comprising:
   allocating a program identity to at least some of the programmes;
   identifying programmes that form episodes of series, and non-series status of programmes that do not form episodes of series;
   allocating episode identities to at least some of such identified episodes; and
   saving the data in a program schedule, said schedule including the program identities of at least some of said programmes, said series and said episode identities of at least some of said programmes that form episodes of series and said non-series status of at least some programmes that do not form episodes of series.

2. A method as claimed in claim 1, wherein said program identity of at least one program is determined by searching for a corresponding program title within the data being processed and on a historical database.

3. A method as claimed in claim 1 or claim 2, wherein said program identity of at least one program is determined by removing characters from the program title and then searching for a corresponding program title within the data being processed and on a historical database.

4. A method as claimed in any one of the preceding claims, wherein programmes that form episodes of series are identified by grouping programmes with similar allocated identities, together.

5. A method as claimed in any one of the preceding claims, wherein programmes that form episodes of a series are identified by searching within the data being processed, for programmes with similar scheduled broadcast times.

6. A method as claimed in any one of the preceding claims, wherein
programmes that form episodes of series are identified by searching within the data being processed and on a historical database, for a series with a title corresponding to a program title.

7. A method as claimed in any one of the preceding claims, wherein programmes that form episodes of a series are identified by assessing the interval between programmes of the same identity.

8. A method as claimed in any one of the preceding claims, which includes determining the frequency with which the episodes of a series are scheduled to be broadcast, by determining the order and times when episodes are scheduled to be broadcast, and assessing from the scheduled broadcast times and order, in accordance with predetermined criteria, whether a scheduled broadcast will be a first broadcast or a repeat of an episode.

9. A method as claimed in claim 8, wherein said predetermined criteria include a ranking of broadcast times.

10. A method as claimed in any one of the preceding claims, which includes reviewing the scheduled broadcast times of programmes with similar identities and identifying the programmes as forming part of a series if the scheduled broadcasts of at least some of the programmes are in close succession in a back-to-back configuration.

11. A method as claimed in any one of the preceding claims, in which the number of programmes is determined, which have similar program identities and which are scheduled to be broadcast within a predetermined time slot, and wherein the programmes are identified as forming episodes of a series if the number of such programmes exceeds a predetermined number.

12. A method as claimed in any one of the preceding claims, in which a series of programmes that is scheduled to be broadcast after a predetermined offset
period during which no other program was broadcast with the same program identity as the programmes of the series, is identified as being the first broadcast of the series.

13. A method as claimed in any one of the preceding claims, which includes using episode identities provided in the unprocessed data, of episodes scheduled to be broadcast before a predetermined cut-off date, to identify the episodes.

14. A method as claimed in any one of the preceding claims, which includes determining whether an episode is being repeated, by searching within the data being processed and on a historical database for episode titles corresponding to the episode identities.

15. A method as claimed in any one of the preceding claims, in which episodes that are being repeated are identified by checking if they are scheduled to be broadcast within a predetermined period of time after the premiere episode.

16. A method as claimed in any one of the preceding claims, in which episodes that are being repeated are automatically identified by the timeslots within which they are scheduled to be broadcast.

17. A method as claimed in any one of the preceding claims, which includes automatically determining a repeat offset by comparing the broadcast times of the premiere of a series with the broadcast times of the first repeat of the series.

18. A method as claimed in claim 17, which includes using said repeat offset to identify episodes of a series chronologically in relation to the premiere of the series.

19. A method as claimed in any one of the preceding claims, which is automatically performed on a computer or other such electronic device.
20. A method as claimed in claim 19, wherein data on said schedule is manually manipulable.

21. A method as claimed in claim 20, wherein data on the schedule is locked to prevent automatic processing of the data.

22. A set of instructions, stored on data storage means and being configured to be executable on a computer or other such electronic device, to perform a method of processing data according to any one of claims 1 to 21.

23. A system for processing data, said system comprising data storage means on which a set of instructions is stored for performing a method of processing data according to any one of claims 1 to 21, and a computer or other such electronic device, configured to perform said method of data processing, when said set of instructions is executed on said computer or other such electronic device.

24. A computer program for performing the method of processing data according to any one of claims 1 to 21.
Fig. 1

1. Supplier
2. Arrange
3. Schedule
4. Enrich
5. Enriched Schedule
6. Database
7. Transmit
8. Subscriber
2/5

Start

Load settings

Loop through all days

Completed

Loop through all programs

Completed

Find/Create Program ID

Loop through channels

Completed

Loop through programs grouped by ID

Completed

Determine series status and frequency

Loop through all days

Completed

Loop through all programs

Completed

Find/Create Episode ID

Completed

Process Guides

Fig. 2
Determine program name

Search database for program name

Not found

Remove unnecessary characters/words from program name

Search database stripped down program name

Not found

Save stripped program name to database

Found

Save program name to database

Found

Assign ID to program

Fig. 3
Fig. 5