

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
16 August 2001 (16.08.2001)

PCT

(10) International Publication Number  
**WO 01/60063 A2**

- (51) International Patent Classification<sup>7</sup>: **H04N 7/00** (74) Agent: **GROENENDAAL, Antonius, W., M.**; Internationaal Octrooibureau B.V., Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).
- (21) International Application Number: PCT/EP01/00463
- (22) International Filing Date: 15 January 2001 (15.01.2001) (81) Designated States (*national*): JP, KR.
- (25) Filing Language: English (84) Designated States (*regional*): European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).
- (26) Publication Language: English
- (30) Priority Data: 09/500,306 8 February 2000 (08.02.2000) US  
**Published:**  
— *without international search report and to be republished upon receipt of that report*
- (71) Applicant: **KONINKLIJKE PHILIPS ELECTRONICS N.V.** [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).  
*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*
- (72) Inventors: **LEE, Kwok, P.**; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). **SCHAFFER, James, D.**; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).



(54) Title: METHOD AND APPARATUS FOR EVALUATING TELEVISION PROGRAM RECOMMENDERS

(57) Abstract: A method and apparatus are disclosed for validating recommendations generated by a television program recommender using programmed viewing agents. A viewing agent is programmed with a set of rules that characterize the viewing preferences of a modeled viewer. During a training phase, the programmed rules of a viewing agent are applied to a set of training programs to obtain an agent viewing history. The generated agent viewing history is processed by a profiler to derive an agent profile containing a set of inferred rules. During an evaluation phase, the programmed rules of the viewing agent are applied to test programs to obtain an agent evaluation viewing set. In parallel, the television program recommender generates a set of program recommendations by applying the agent profile to the test programs. The agent evaluation viewing set (generated from the programmed rules) is then compared with the program recommendations (generated from the inferred rules). The present invention also estimates the required size of the viewing history to provide a given level of accuracy. The viewing agents can be programmed to introduce one or more random shows into the viewing history, or to change the preferences of the viewing agent over time, thereby allowing the television program recommender validator to evaluate how the program recommender processes such non-stationaries.

WO 01/60063 A2

## Method and apparatus for evaluating television program recommenders

### Field of the Invention

The present invention relates to methods and apparatus for recommending television programming, and more particularly, to techniques for evaluating television program recommenders.

5

### Background of the Invention

As the number of channels available to television viewers has increased, along with the diversity of the programming content available on such channels, it has become increasingly challenging for television viewers to identify television programs of interest.

10 Historically, television viewers identified television programs of interest by analyzing printed television program guides. Typically, such printed television program guides contained grids listing the available television programs by time and date, channel and title. As the number of television programs has increased, it has become increasingly difficult to effectively identify desirable television programs using such printed guides.

15 More recently, television program guides have become available in an electronic format, often referred to as electronic program guides (EPGs). Like printed television program guides, EPGs contain grids listing the available television programs by time and date, channel and title. Some EPGs, however, allow television viewers to sort or search the available television programs in accordance with personalized preferences. In  
20 addition, EPGs allow for on-screen presentation of the available television programs.

While EPGs allow viewers to identify desirable programs more efficiently than conventional printed guides, they suffer from a number of limitations, which if overcome, could further enhance the ability of viewers to identify desirable programs. For example, many viewers have a particular preference towards, or bias against, certain  
25 categories of programming, such as action-based programs or sports programming. Thus, the viewer preferences can be applied to the EPG to obtain a set of recommended programs that may be of interest to a particular viewer.

Thus, a number of tools have been proposed or suggested for recommending television programming. The Tivo™ system, for example, commercially available from Tivo,

Inc., of Sunnyvale, California, allows viewers to rate shows using a "Thumbs Up and Thumbs Down" feature and thereby indicate programs that the viewer likes and dislikes, respectively. Thereafter, the TiVo receiver matches the recorded viewer preferences with received program data, such as an EPG, to make recommendations tailored to each viewer.

5 There is currently no way, however, to validate the recommendations generated by such tools for recommending television programming, short of testing the tools with human volunteers. In addition, there is no way to determine when the recommendations have achieved a given level of accuracy. A need therefore exists for a method and apparatus for validating  
10 recommendations generated by a television program recommender. A further need exists for a method and apparatus for determining when the size of the viewing history (record of shows watched/not watched) is sufficient to provide a given level of accuracy.

#### Summary of the Invention

15 Generally, a method and apparatus are disclosed for evaluating the effectiveness of a television program recommender by evaluating program recommendations generated by the television program recommender for one or more programmed viewing agents. A viewing agent is programmed with a set of rules that characterize the viewing preferences of a modeled viewer. Viewing agents of varying complexity and having varying  
20 program preferences can be constructed by defining various rules that characterize program attributes, such as genre, actors and program duration.

During a training phase of the present invention, a viewing agent containing the programmed rules is applied to a set of training programs from an electronic program guide (EPG) to algorithmically obtain an agent viewing history, indicating whether the  
25 viewing agent would have liked or disliked each training program. The generated agent viewing history is then processed by the profiler portion of the television program recommender being evaluated. The profiler derives an agent profile containing a set of inferred rules that attempt to mimic the programmed preferences of the viewing agent. Thus, the profiler attempts to derive the viewing habits of the viewing agent based on the set of  
30 programs that the viewing agent liked or disliked.

During an evaluation phase of the present invention, the programmed rules of the viewing agent are applied to test data from an electronic program guide (EPG) to obtain an agent evaluation viewing set. In parallel, the television program recommender generates a set of program recommendations by applying the agent profile generated during the training

phase to the test data from the electronic program guide (EPG). The present invention can then compare the agent evaluation viewing set (generated from the programmed rules) with a set of program recommendations produced by the television program recommender being evaluated (generated from the inferred rules). In this manner, the effectiveness of the television program recommender can be evaluated.

According to another aspect of the invention, a method and apparatus are disclosed for determining the required size of the viewing history to provide a given level of accuracy. Generally, the size of the required viewing history can be obtained by utilizing viewing histories of varying sizes for the same viewing agent. Thereafter, the size of the viewing history required to exceed a predefined accuracy threshold can be obtained.

The viewing agents can be programmed to introduce one or more random shows into the viewing history, allowing the television program recommender validator to determine how the program recommender processes such noise. In a further variation, the viewing agents can be programmed to gradually change the programmed viewer preferences over time, allowing the television program recommender validator to determine how the program recommender responds to such changes.

A more complete understanding of the present invention, as well as further features and advantages of the present invention, will be obtained by reference to the following detailed description and drawings.

#### Brief Description of the Drawings

Fig. 1A illustrates a television program recommender validator in accordance with a training phase of the present invention;

Fig. 1B illustrates a television program recommender validator in accordance with an evaluation phase of the present invention;

Fig. 2 is a schematic block diagram of the television program recommender validator of Fig. 1B;

Fig. 3 is a flow chart describing the validation process (training phase) of Fig. 2 embodying principles of the present invention;

Fig. 4 is a flow chart describing the validation process (evaluation phase) of Fig. 2 embodying principles of the present invention;

Fig. 5 is a flow chart describing the sensitivity analysis process of Fig. 2 embodying principles of the present invention;

Fig. 6A illustrates a confusion matrix that analyzes whether a viewing agent would have liked a given show according to the programmed preferences and whether the program recommender actually recommended the show;

Fig. 6B illustrates a plot of the precision and recall values of the training data as a function of the number of shows viewed (size of viewing history); and

Fig. 6C illustrates a plot of the number of shows viewed varying as a function of the complexity of the viewing agent.

### **Detailed Description**

10 Figs. 1A and 1B illustrate a television program recommender validator 200, discussed further below in conjunction with FIG. 2, in accordance with a training phase and evaluation phase, respectively, of the present invention. According to one aspect of the present invention, the television program recommender validator 200 evaluates the effectiveness of a television program recommender 160 by evaluating program

15 recommendations generated by the television program recommender 160 for one or more programmed viewing agents 120 (or algorithmic viewers). As used herein, a viewing agent 120 is any programmed viewer whose viewing habits are predetermined. Each viewing agent 120 is programmed with a set of rules that categorize the viewing preferences of the modeled viewer.

20 Viewing agents 120 of varying complexity and having varying program preferences can be constructed in accordance with the present invention by defining various rules that characterize programs by their attributes. For example, a viewing agent 120 may be programmed with a strong preference for programs that belong to one or more particular genres, such as the "action" genre. In further variations, viewing agents 120 may be

25 programmed with a preference for shows containing one or more identified actors and having a given duration.

For example, if a viewing agent 120 is programmed with a strong preference for programs that belong to the action genre, all action shows should be identified by the program recommender 160. The programs identified directly from the programmed rules of

30 the viewing agent 120 in accordance with the present invention can be compared with the program recommendations 130 produced by the television program recommender 160 utilizing inferred rules contained in an agent profile 140. In this manner, the effectiveness of the television program recommender 160 can be evaluated.

According to another aspect of the invention, the television program recommender validator 200 estimates the size of the viewing history needed for the television program recommender 160 to provide results with a predefined reliability. Generally, the size of the required viewing history can be obtained by utilizing viewing histories of varying sizes for the same viewing agent 120. Thereafter, the size of the viewing history required to exceed a predefined accuracy can be obtained.

The present invention can evaluate any television program recommender 160, such as the television program recommenders 160 described in United States Patent Application Serial No. 09/466.406, filed 17.12.1999, entitled "Method and Apparatus for Recommending Television Programming Using Decision Trees," (Attorney Docket No. 700772) and United States Patent Application Serial No. 09/498.271, filed 04.02.2000, entitled "Bayesian TV Show Recommender," (Attorney Docket No. 700690), or the Tivo™ system, commercially available from Tivo, Inc., of Sunnyvale, California. It is noted that the profiler 130 is generally a component of the television program recommender 160.

15

#### TRAINING PHASE

Fig. 1A illustrates a training phase of the present invention. The present invention applies the viewing agent 120 containing the programmed rules to a set of training programs from an electronic program guide (EPG) 110 to algorithmically obtain an agent viewing history 125. The agent viewing history 125 indicates whether the viewing agent 120 would have liked or disliked each of the training programs 110.

As shown in Fig. 1A, the profiler portion 130 of the television program recommender 160 (Fig. 1B) being evaluated then processes the agent viewing history 125 to derive an agent profile 140 containing a set of inferred rules that mimic the programmed preferences of the viewing agent 120. Thus, the profiler 130 attempts to derive the viewing habits of the viewing agent 120 based on the set of programs that the viewing agent 120 liked or disliked. Generally, more accurate program recommenders 160 will generate sets of inferred rules that more closely resemble the programmed rules of the viewing agent 120.

#### EVALUATION PHASE

Fig. 1B illustrates an evaluation phase of the present invention. Generally, the evaluation phase applies the programmed rules of the viewing agent 120 to a set of test data 150 from an electronic program guide (EPG) to obtain an agent evaluation viewing set 170. The television program recommender validator 200 then compares the agent evaluation

viewing set 170 with a set of program recommendations 180 produced by the television program recommender 160 being evaluated. The television program recommender 160 generates the program recommendations 180 by applying the agent profile 140 generated during the training phase to the test data 150 from an electronic program guide (EPG).

5 Fig. 2 is a schematic block diagram of the television program recommender validator 200. The television program recommender validator 200 may be embodied as any computing device, such as a personal computer or workstation, that contains a processor 210, such as a central processing unit (CPU), and memory 220, such as RAM and ROM. As shown in Fig. 2, the program recommender validator 200 contains a validation process  
10 (training phase) 300, discussed further below in conjunction with Fig. 3, a validation process (evaluation phase) 400, discussed further below in conjunction with Fig. 4, and a sensitivity analysis process 500, discussed further below in conjunction with Fig. 5.

Generally, the validation process (training phase) 300 implements the training phase, discussed above, to generate the agent profile 140. The validation process (evaluation  
15 phase) 400 implements the evaluation phase, discussed above, to compare the agent evaluation viewing set 170 generated directly from the programmed rules for the viewing agent 120 with a set of program recommendations 180 produced by the television program recommender 160 being evaluated. The sensitivity analysis process 500 estimates the size of the viewing history needed for the television program recommender 160 to provide results  
20 with a predefined reliability.

Fig. 3 is a flow chart describing the validation process (training phase) 300 embodying principles of the present invention. The validation process (training phase) 300 implements the training phase, discussed above, to generate the agent profile 140. As shown in Fig. 3, the validation process (training phase) 300 initially constructs a viewing agent 120  
25 by defining various rules that characterize programs by their attributes as being liked or disliked during step 310. The test data from the electronic program guide (EPG) 110 is then applied to the programmed rules of the viewing agent 120 during step 320 to generate the agent viewing history 125.

Finally, the profiler 130 is applied to the agent viewing history 125 during step  
30 330 to identify the inferred rules that characterize the agent viewing history 125. Program control then terminates.

Fig. 4 is a flow chart describing the validation process (evaluation phase) 400 embodying principles of the present invention. The validation process (evaluation phase) 400 implements the evaluation phase, discussed above, to compare the agent evaluation

viewing set 170 generated directly from the programmed rules for the viewing agent 120 with a set of program recommendations 180 produced by the television program recommender 160 being evaluated.

As shown in Fig. 4, the validation process (evaluation phase) 400 initially  
5 applies the test data from the EPG 150 to the programmed rules of the viewing agent 120 during step 410 to generate the agent evaluation viewing set 170 indicating programs that the programmed viewing agent 120 would watch. Thereafter, the validation process (evaluation phase) 400 applies the test data from the EPG 150 to the TV recommender 160, using the agent profile 140 generated during the training phase, to generate a set of program  
10 recommendations 180 during step 420.

During step 430, the validation process (evaluation phase) 400 then compares the set of programs 170 identified during step 410 from the program rules with the set of program recommendations 180 identified by the program recommender 160 during step 420 from the inferred rules.

15 Finally, the effectiveness of the television program recommender 160 can be evaluated during step 440, based on the comparison. For example, the effectiveness of the television program recommender 160 can be evaluated using well-known pattern recognition techniques, such as the mean square error, precision and recall (discussed below) or receiver operator characteristic curves. Program control then terminates.

20 Fig. 5 is a flow chart describing the sensitivity analysis process 500 embodying principles of the present invention. Generally, the sensitivity analysis process 500 estimates the size of the viewing history needed for the television program recommender 160 to provide results with a predefined reliability or accuracy.

As shown in Fig. 5, the sensitivity analysis process 300 initially generates  
25 viewer agents of varying complexity during step 510. For each viewing agent 120, the sensitivity analysis process 500 then generates viewing histories of varying size during step 520. The validation process (evaluation) 400 is then executed during step 525 for each viewing history size and viewing agent 120 to compare the agent evaluation viewing set 170 generated directly from the programmed rules for the viewing agent 120 with a set of  
30 program recommendations 180 produced by the television program recommender 160 being evaluated.

Thereafter, for each viewing history size and viewing agent 120, the sensitivity analysis process 500 computes the precision of the recommendations during step 530 as follows:



$$\text{Precision} = \frac{TP}{CT} = \frac{TP}{TP+FP},$$

where TP indicates the true positives, FP indicates the false positives, and the column total, CT, is equal to the TP plus the FP.

In addition, for each viewing history size and viewing agent 120, the  
 5 sensitivity analysis process 500 computes the recall of the training data during step 540 as follows:

$$\text{Recall} = \frac{TP}{RT} = \frac{TP}{TP+FN},$$

where FN indicates the false negatives and the row total, RT, is equal to TP plus FN.

The concepts and calculations for precision and recall are discussed further below in  
 10 conjunction with FIG. 6A. The required size of the viewing history for the precision and recall to exceed a predefined threshold is then determined during step 550 and the size of viewing history is plotted as a function of the complexity of the viewing agent 120 during step 560. Program control then terminates.

## 15 SENSITIVITY ANALYSIS

Fig. 6A illustrates a well-known confusion matrix 600 that analyzes whether the viewing agent 120 would have liked a given show according to the programmed preferences (satisfied rules in profile), and whether the program recommender 160 actually recommended the show. The confusion matrix 600 indicates the true positives (TP), false positives (FP), true  
 20 negatives (TN) and false negatives (FN). The confusion matrix 600 also indicates the row total (RT = TP + FN) and the column total (CT = TP + FP).

In accordance with well-known pattern recognition techniques, such as those described in Boreczky and Rowe, "Comparison of Video Shot Boundary Detection Techniques," SPIE, Vol. 2670, 170-79 (1996), incorporated by reference herein, the precision  
 25 and recall of the training set can be evaluated according to the following expressions:

$$\text{Precision} = \frac{TP}{CT} = \frac{TP}{TP+FP}$$

$$\text{Recall} = \frac{TP}{RT} = \frac{TP}{TP+FN}$$

Fig. 6B illustrates a plot 650 of the precision and recall values as a function of  
 30 the number of shows viewed (size of the viewing history). The precision and recall values

vary between values of zero and one, and generally increase as the size of the viewing history increases. To obtain a desired degree of accuracy in accordance with the present invention, the size of the viewing history can be adjusted such that both the precision and recall values exceed a predefined threshold.

5                    Fig. 6C illustrates a plot 680 of the size of the viewing history needed to achieve a given level of accuracy or reliability, as a function of the complexity of the viewing agent 120. Clearly, the number of shows that must be included in the viewing history must be increased to achieve a given accuracy level as the complexity of the viewing agent 120 increases.

10                    In one variation, the viewing agents 120 can be programmed to introduce one or more random shows into the viewing history 125 and thereby determine how the program recommender 160 processes such noise. In a further variation, the viewing agents 120 can be programmed to gradually change the programmed viewer preferences over time, and thereby determine how the program recommender 160 responds to such changes.

15                    It is to be understood that the embodiments and variations shown and described herein are merely illustrative of the principles of this invention and that various modifications may be implemented by those skilled in the art without departing from the scope and spirit of the invention.

## CLAIMS:

1. A method for validating program recommendations (180) produced by a program recommender (160), comprising the steps of:
  - generating a viewing agent (120) using one or more programmed rules that characterize viewing preferences;
  - 5 applying said programmed rules to a set of training programs (110) to obtain an agent viewing history (125);
  - processing said agent viewing history (125) with a profiler (130) to derive an agent profile (140) containing a set of inferred rules;
  - applying said programmed rules to a set of test programs (150) to obtain an
  - 10 agent evaluation viewing set (170);
  - generating said program recommendations (180) produced by said program recommender (160) by applying said agent profile (140) to said test programs (150); and
  - comparing said agent evaluation viewing set (170) with said program recommendations (180).
- 15 2. The method of claim 1, wherein said viewing agent (120) is selected to have viewing habits that model the viewing habits of a type of viewer.
3. The method of claim 1, wherein said agent viewing history (125) contains an
- 20 indication of whether said viewing agent (120) liked or disliked each program in a set of training programs (110).
4. The method of claim 1, wherein said agent evaluation viewing set (170) contains an indication of whether said viewing agent (120) liked or disliked each program in
- 25 a set of test programs (150).
5. The method of claim 1, wherein said viewing preferences characterize programs by their attributes that are liked or disliked by said viewing agent (120).

6. The method of claim 1, wherein said profiler (130) is a component of said program recommender (160).

7. The method of claim 1, wherein said inferred rules mimic said programmed  
5 preferences of the viewing agent (120).

8. The method of claim 1, wherein said viewing preferences include one or more random programs.

10 9. The method of claim 1, wherein said viewing preferences change over time.

10. The method of claim 1, wherein said comparing step further comprises the step of comparing calculated precision and recall values to a predefined level of accuracy.

15 11. A method for validating program recommendations (180) produced by a program recommender (160) using a viewing agent (120) having programmed viewing preferences, said method comprising the steps of:

processing a viewing history (125) of said viewing agent (120) using a profiler (130) to generate an agent profile (140), said agent profile (140) containing a set of inferred  
20 rules that characterize said programmed viewing preferences;

applying said programmed viewing preferences to a set of test programs (150) to obtain an agent evaluation viewing set (170);

generating said program recommendations (180) produced by said program recommender (160) by applying said agent profile (140) to said test programs (150); and

25 comparing said agent evaluation viewing set (170) with said program recommendations (180).

12. A method for determining the required size of a viewing history (125) for a program recommender (160) to provide a given level of accuracy for a user, said method  
30 comprising the steps of:

generating a viewing agent (120) using one or more programmed rules that characterize viewing preferences;

generating a plurality of viewing histories (125) of varying sizes for said viewing agent (120);

determining the precision for each viewing history;  
determining the recall for each viewing history; and  
determining a required size for said viewing history (125) such that said  
precision and recall values exceed a predefined threshold.

5

13. The method of claim 12, wherein said precision is determined from the true  
positives (TP) and false positives (FP) as follows:

$$\text{Precision} = \frac{TP}{CT} = \frac{TP}{TP+FP}$$

10 14. The method of claim 12, wherein said precision is determined from the true  
positives (TP) and false negatives (FN) as follows:

$$\text{Recall} = \frac{TP}{RT} = \frac{TP}{TP+FN}$$

15 15. A system (200) for validating program recommendations (180) produced by a  
program recommender (160), comprising:

a memory (220) for storing computer readable code; and  
a processor (210) operatively coupled to said memory (220), said processor  
(210) configured to:

20 generate a viewing agent (120) using one or more programmed rules that  
characterize viewing preferences;

apply said programmed rules to a set of training programs (110) to obtain an  
agent viewing history (125);

process said agent viewing history (125) with a profiler (130) to derive an  
agent profile (140) containing a set of inferred rules;

25 apply said programmed rules to a set of test programs (150) to obtain an agent  
evaluation viewing set (170);

generate said program recommendations (180) produced by said program  
recommender (160) by applying said agent profile (140) to said test programs (150); and  
compare said agent evaluation viewing set (170) with said program recommendations (180).

30

16. A system (200) for validating program recommendations (180) produced by a program recommender (160) using a viewing agent (120) having programmed viewing preferences, comprising:

a memory (220) for storing computer readable code; and

5 a processor (210) operatively coupled to said memory (220), said processor (210) configured to:

process a viewing history (125) of said viewing agent (120) using a profiler (130) to generate an agent profile (140), said agent profile (140) containing a set of inferred rules that characterize said programmed viewing preferences;

10 apply said programmed viewing preferences to a set of test programs (150) to obtain an agent evaluation viewing set (170);

generate said program recommendations (180) produced by said program recommender (160) by applying said agent profile (140) to said test programs (150); and compare said agent evaluation viewing set (170) with said program recommendations (180).

15

17. A system (200) for determining the required size of a viewing history for a program recommender (160) to provide a given level of accuracy for a user, comprising:

a memory (220) for storing computer readable code; and

20 (210) configured to:

generate a viewing agent (120) using one or more programmed rules that characterize viewing preferences;

generate a plurality of viewing histories (125) of varying sizes for said viewing agent (120);

25 determine the precision for each viewing history;

determine the recall for each viewing history; and

determine a required size for said viewing history such that said precision and recall values exceed a predefined threshold.

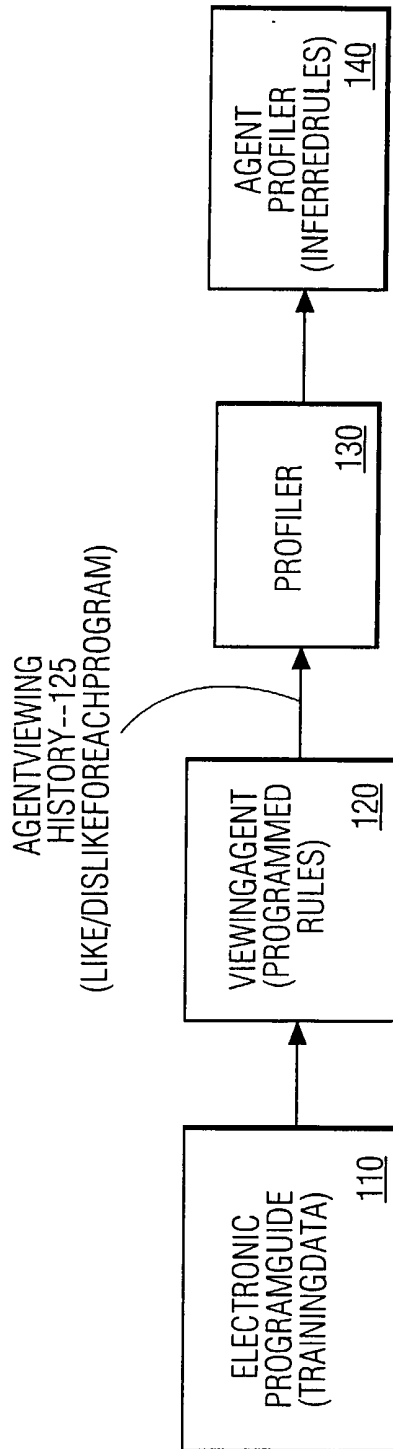


FIG. 1A

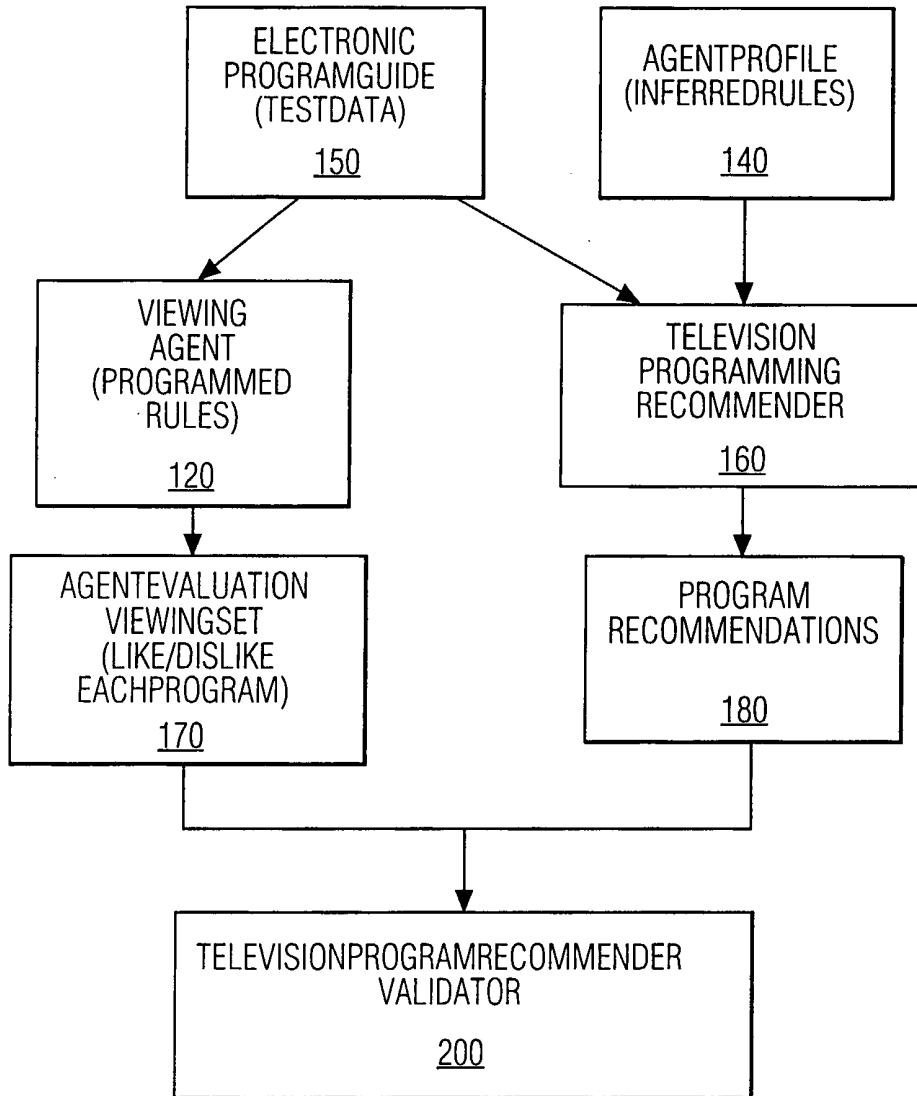


FIG. 1B



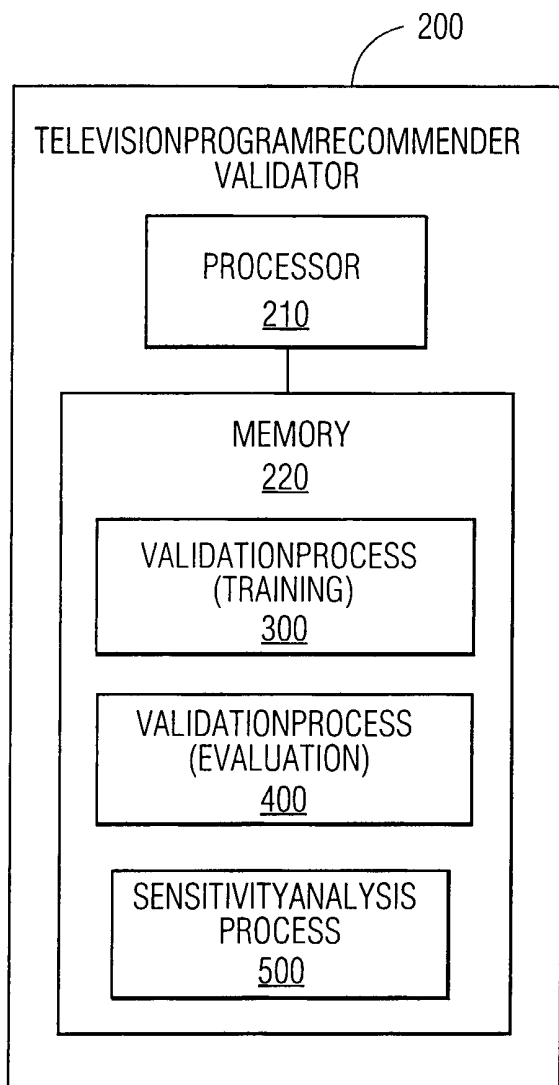


FIG. 2

4/7

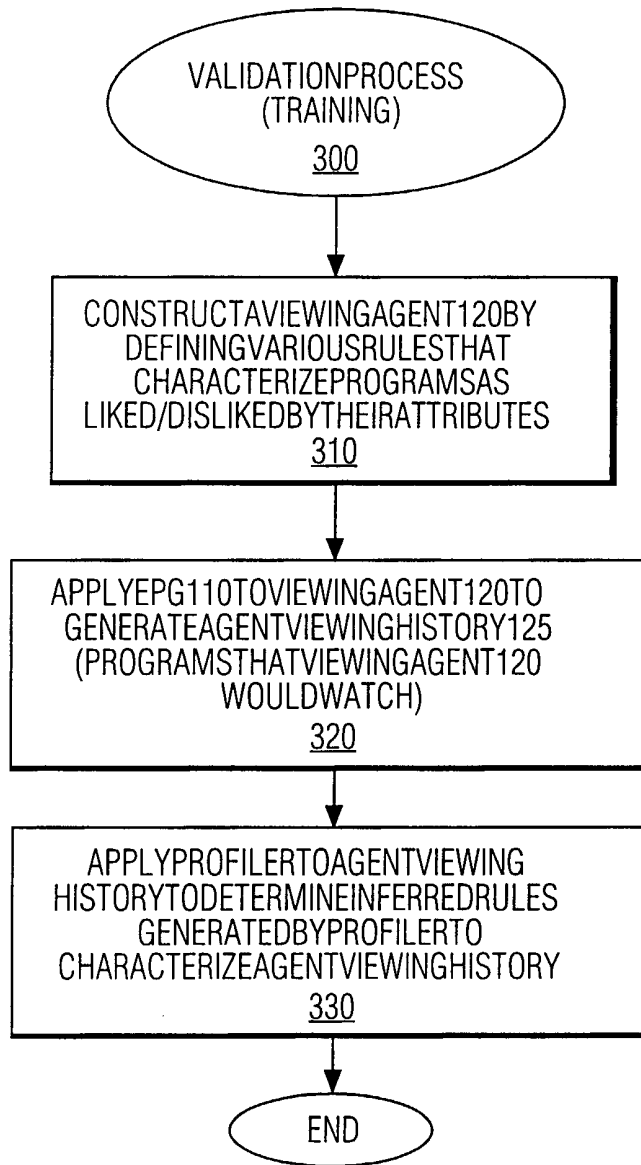


FIG. 3

5/7

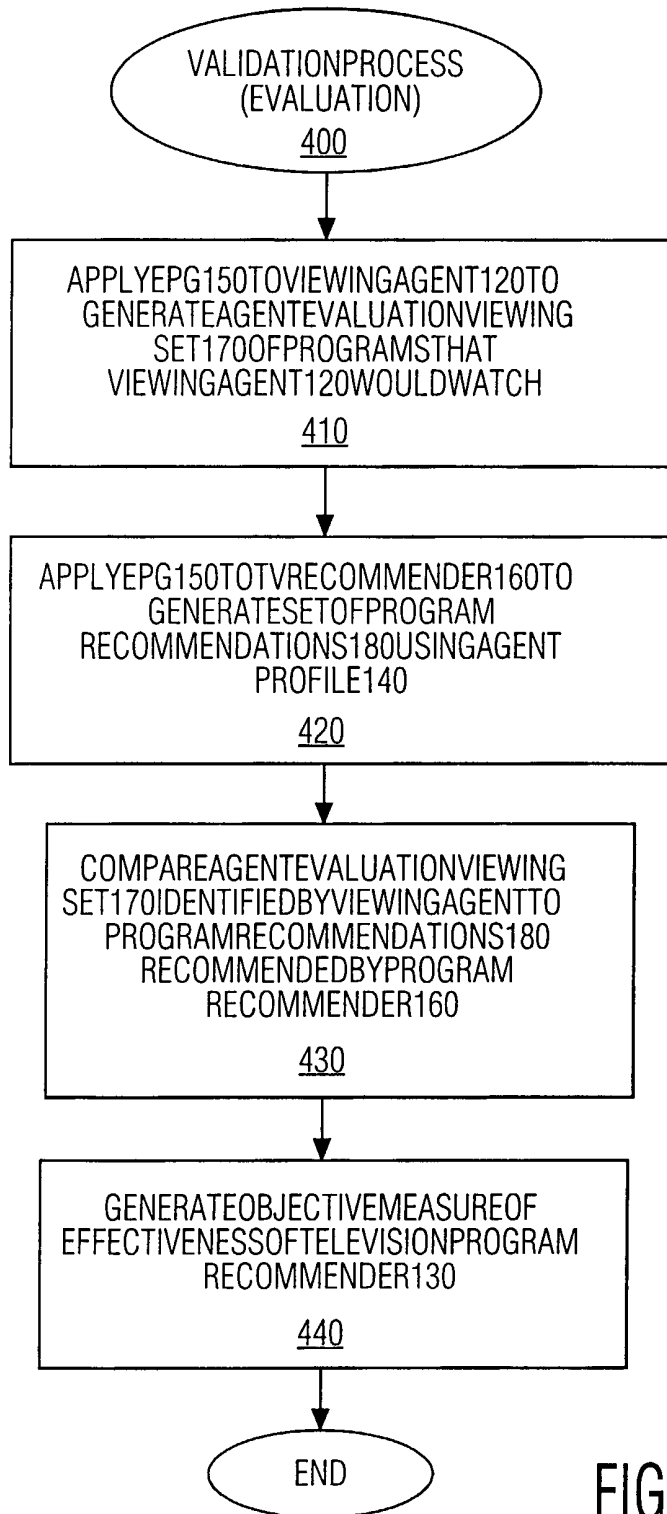


FIG. 4

6/7

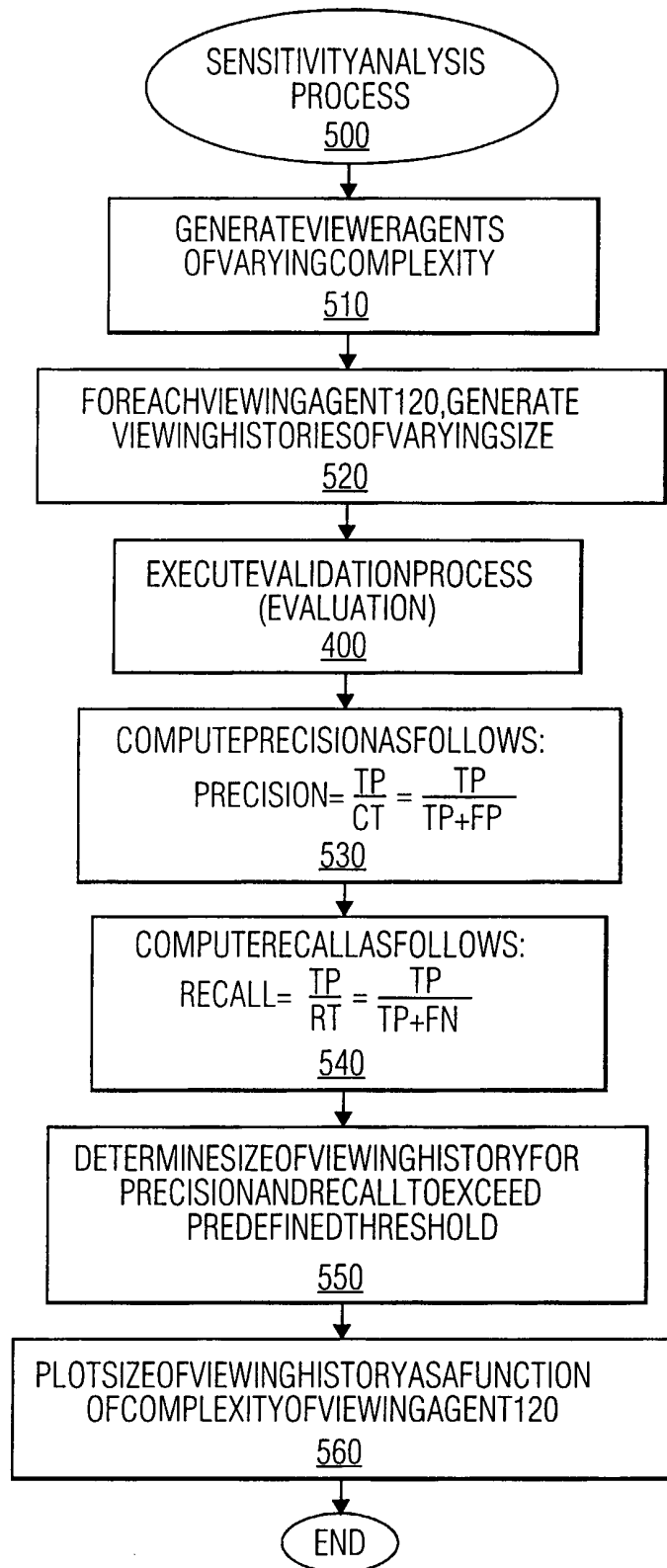


FIG. 5

RECOMM			
	TP	FN	RT
TRUE			
	FP	TN	
	CT		

FIG. 6A

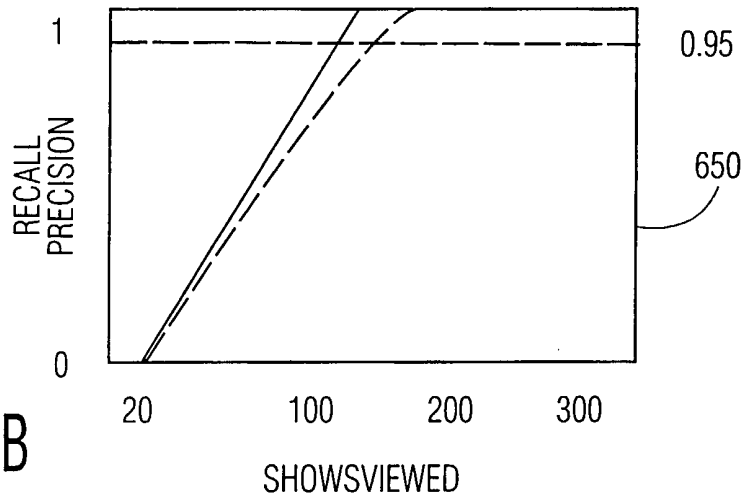


FIG. 6B

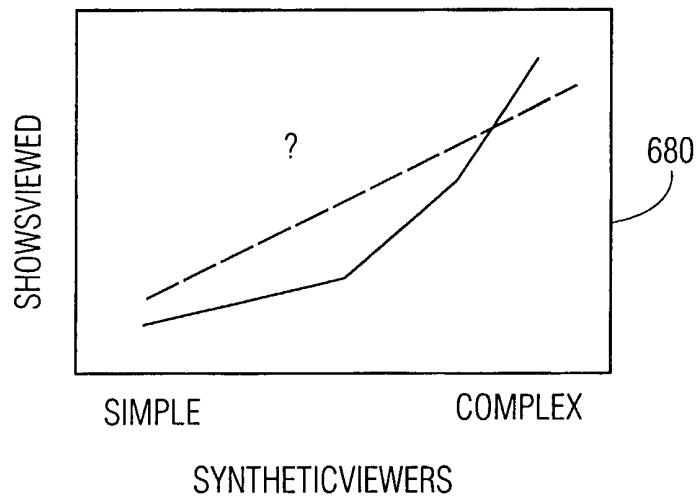


FIG. 6C