(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 November 2012 (15.11.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/152513 A2

(51) International Patent Classification:
G06F 17/30 (2006.01) *G06Q 10/00* (2012.01)
H04L 29/06 (2006.01)

(21) International Application Number:
PCT/EP2012/056086

(22) International Filing Date:
3 April 2012 (03.04.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11305534.7 6 May 2011 (06.05.2011) EP

(71) Applicant *(for all designated States except US)*: **ALCATEL LUCENT** [FR/FR]; 3 avenue Octave Gréard, F-75007 Paris (FR).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **BENOIT, Christophe** [FR/FR]; Alcatel Lucent Bell Labs France, Centre de Villarceaux, F-91620 Nozay (FR). **VERDOT, Vincent** [FR/FR]; Alcatel Lucent Bell Labs France, Centre de Villarceaux, F-91620 Nozay (FR).

(74) Agent: **MOUNEY, Jérome**; Alcatel-Lucent International, 32 avenue Kléber, F-92700 Colombes (FR).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with declaration under Article 17(2)(a); without abstract; title not checked by the International Searching Authority*

WO 2012/152513 A2

(54) Title: METHOD AND SYSTEM FOR PREDICTING SEARCH TYPES FOR THE INTERNET OF THINGS

(57) Abstract:

METHOD AND SYSTEM FOR PREDICTING SEARCH TYPES FOR THE INTERNET OF THINGS

The present invention pertains to the technical field of the Internet of things, and more specifically to searching for connected devices in intelligent spaces.

The "Internet of things" means various technical solutions (wireless technologies, Web technologies, computer technologies, TCP/IP, and electronic identification technologies, for example), making it possible to extend the Internet to include the objects that surround us. We can then speak of the Internet of things or, by transposing Web standards, the Web of things.

The Internet of things, interconnecting objects ranging from simple objects identified by electronic labels to devices equipped with actively communicating computer systems, makes it possible to identify these objects, to receive, store, process, and transfer data related to them between/in the physical and virtual environments. These environments are referred to as intelligent spaces.

Here, "intelligent space" means an environment comprising a plurality of sensors, effectors, fluid means of communication (ubiquitous/distributed computing) and on-board intelligence (environmental intelligence) arranged to guarantee information management in a variety of contexts.

As an example of these environments, one may cite:

- A professional environment (an office, a meeting room, an open space, or a company department for example) comprising a plurality of devices such as fixed/mobile telephones, GPSs, desktop/laptop computers, printers, photocopiers, fax machines, cameras, screens, video projectors, external hard disks, or paper files with RFID labels for example;
- A home environment (a bedroom, a house, a hotel, or a residence for example) comprising a plurality of devices such as home appliance systems (alarms, interphones, Smartphones, computers, heating systems, remote controls, televisions, radio receivers,

stereo systems, home appliances such as washing machines, freezers, electric cookers,) or manufactured products (labelled lamps, yogurt containers, bottles, or chairs for example);

- A medical, industrial, transport, or commercial environment for example, such as an airport, a train station, a shopping centre, a shopping street, a sports complex, or various buildings at a company.

Carried by continuous technological progress (wireless network communication, the development of mobility, ubiquitous computing, the diversification of identification and connection systems) and a profound change in society (autonomous socio-technical environments focused on the individual, personal networks or PANs for Personal Area Networks), the Internet of things is continuously incorporating an increasing number of devices used on a daily basis. The document ("More than 50 Billion connected devices", Ericsson white paper, 284 23-3149 Uen, February 2011), estimates that in the near future more than 50 billion connected devices will be covered by the Internet of things.

Now the problem raised is how to accelerate/improve searching for a specific device from among a very large number of connected devices. In fact, faced with this growing mass of connected devices, the conventional search engines of the Internet of things encounter great difficulty in finding an adequate compromise between accuracy and speed for carrying out a search.

To respond to a search query on the Internet of things, an algorithm browsing a graph of descriptions of connected devices, or a method based on an automatic learning technique, can be adopted to return results above a certain threshold.

However, these search methods can only be implemented when a search context is explicitly indicated in the search query. As a search context is not necessarily known a priori, it is therefore typically difficult to identify the appropriate search type: a search privileging speed or accuracy?

One object of the present invention is to propose a framework that makes it possible to predict a search context on the Internet of things.

Another object of the present invention is to predict the type of a search (precise, fuzzy) to be carried out in response to a search query on the Internet of things.

Another object of the present invention is to improve the resolution of search queries on the Internet of things.

Another object of the present invention is to reduce the latency of search engines by predicting a search context.

One object of the present invention is to make a substantial improvement to searching (speed/accuracy) on the Internet of things.

Another object of the present invention is to improve search results in a very widely distributed and open information system such as the Internet of things.

For this purpose, the invention relates, according to a first aspect, to a method for predicting the search type for a search query on the Internet of things comprising a step to estimate the distribution of the desired object using a means of calculation.

This method also comprises a step to calculate search type probabilities based on the search types of previous search queries with the same identifier as the search query.

This method also comprises a step for weighting the probabilities calculated by weights according to the estimated distributions (that is to say, according to the organisation of an intelligent space: large number of connected devices of various types for example).

The invention relates, according to a second aspect, to a system for predicting the search type in an intelligent space comprising means adapted for studying the distribution of the types of connected devices present in the intelligent space.

This system also comprises a context analyser adapted for extracting, from a search query, a search identifier, and the type of device sought.

According to a third aspect, the invention proposes a computer program product implemented on a memory medium, which may be implemented within an information processing unit, and comprises instructions for incrementing the method summarized above.

Other characteristics and advantages of the invention will become more clearly and completely apparent upon reading the description below of preferred embodiments, which is done with reference to

- Figure 1, which schematically depicts the application context of one embodiment; and
- Figure 2, which schematically depicts a functional, non-limiting depiction of one embodiment.

Figure 1 shows an intelligent space **1** covering a plurality of connected devices **21-37** comprised by the Internet to things. The intelligent space **1** is generally the result of the cooperation and interconnection of a plurality of intelligent sub-spaces **2-4** and/or connected devices **24-26, 32, 37**.

A distributed system **5** for information processing (comprising ONSs (Object Naming Services), ubiquitous communication networks, computer hardware/equipment means, and programmed sensors for example) deployed cooperatively in the intelligent space **1** is adapted for supporting the Internet of things in this space. This makes it possible for a user, via a Web interface, to interact (verify the presence, display information, or use a functionality of a connected device, for example) with the connected devices **21-37**. For example, this may consist of

- Remotely turning off a lamp equipped with an electronic identification system (an RFID label or a contactless chip for example);
- Starting recording of a film showing on a television channel;
- Verifying the presence of a connected device in a warehouse;
- Recovering information (merchandise transit, validity expiration date, or origin for example) of a connected product.

In particular, the distributed information processing system **5** comprises search means (one or more search engines) configured to carry out a search on the Internet of things in response to a search query.

Preferably, the intelligent sub-spaces **2-6** are equipped, respectively, with a local search engine, the local search engines being federated by a global search engine.

A search on the Internet of things may also be obtained by implementing distributed solutions that make it possible to support the handling of a large number of search queries.

A search query may be initiated by a user or an application in the space or an intelligent sub-space, through a connected device **21-37** (a computer, a Smartphone, or a PDA for example).

When a local search engine receives a search query from a user or an application, a search type prediction system then establishes a search context before carrying out this search.

The "search context" means the search type or mode to be carried out. The main search contexts that can be distinguished are:

- A fuzzy search based on the terms of the query or variations of them (by using for example an edit distance such as the Jaccard distance to define the concept of proximity);
- A precise search based on a precise query syntax.

Figure 2 depicts the various steps that make it possible to predict a search context for each search query received by the search type prediction system. It should be noted that these steps are designed to define a search context and not the search itself, which is the responsibility of the search engine.

To do that, the search type prediction system comprises a context analyser **10**.

When received, a search query is analysed by a context analyser **10** to extract

- At least one identifier (query_ID) (arrow **101** in Figure 2). This identifier may be an IP address, a login, a cookie identifier, or any

other data that makes it more or less possible to uniquely distinguish the initiator of the query.

- The type of the device sought by this query (arrow **108** in Figure 2).

The search type prediction system checks to see if this query_ID identifier has a search profile in an internal database **11**. In other words, the search type prediction system searches to see whether or not the initiator of the query received has a history of queries from the internal database **11**.

To that end, the search type prediction system uses the query_ID identifier to recover previous search types corresponding to this identifier from the internal database **11**. It should be noted here that the previous search types are recovered, not the entire query/result history associated with the query identifier "query_ID". These data are recovered to calculate statistics on the previous search types (that is to say, created in response to the previous search queries sent by the initiator of the received query), and to calculate search type probabilities from them for the received query.

A test **101** is carried out on the history corresponding to the identifier query_ID found in the internal database **11** to check whether or not it is empty.

If the history is empty (arrow **112** in Figure 2), then the prediction system queries a global discovery module **123**. This module **123** comprises a global context analyser **12** that is connected to a plurality of local context analysers **13**.

Preferably, the global context analyser **12** queries a local context analyser **13** for an intelligent sub-space that is spatially near the context analyser **10** that received the search query. In other words, if there is no history local to the context analyser **10**, the prediction system will use the intelligent sub-spaces located near the initiator of the query.

An intelligent sub-space near the initiator of a search query may be found using information extracted from the received search query (an IP address or GPS coordinates for example).

Advantageously, the intelligent sub-spaces close to each other will generally comprise similar connected devices (clusters of devices) constituting a sort of various "interest centres" for the initiator of a query.

The non-empty history recovered from the internal database **11** (arrow **111** in Figure 2) or via the global discovery module **123** (arrow **113** in Figure 2) and corresponding to the identifier query_ID is ranked (step **14** in Figure 2) in a table displaying statistics (total number, number of searches by type) of previous search types as well as their respective probabilities. As an example, this table may comprise the following data:

Total searches: 50

Fuzzy searches: 45

Precise searches: 5

Probability(fuzzy searches): 0.9

Probability(precise searches): 0.1

Additionally, the search type prediction system **52** determines the distribution of the type of device sought. A semantic approach may be adopted in particular to identify the type of object requested (arrow **108** in Figure 2) in a search query. Distribution means the organisation of connected devices in an intelligent space: a large number of connected devices, varying types of connected devices for example.

To do this, a connected device analyser **18** makes it possible to study (provide information about) the distribution of the types of connected devices present in the intelligent space **1**, such as for example:

- the intelligent space **1** is comprised of a plurality of connected devices of various types or of the same type (a plurality of lamps, a plurality of telephones, for example);
- The connected devices may be spread into several clusters;
- The types of connected devices are highly heterogeneous;

The connected device analyser **18** aims to provide a general picture of the connected devices located in the intelligent space **1** in terms of "distribution".

The test **180** of the distribution of the types of connected devices makes it possible to estimate (approach) using means of calculation, the distribution of the type of object sought with a theoretical distribution. As examples, this distribution may be

- A Gaussian distribution **182** (that is to say few clusters of connected devices) that favours (assigns a significant weight to) a precise search over a fuzzy search (step **20** in Figure 2);
- A stochastic distribution **181** (that is to say, a large number of connected devices) that favours (assigns a significant weight to) a fuzzy search over a precise search (step **19** in Figure 2).

The weights assigned to search types (steps **19** and **20**) according to the distributions of the type of device sought on the Internet of things are used to refine (step **15** in Figure 2) the search type probabilities calculated in step **14**.

As an example, in step **17**, the search type probabilities calculated in step **14** are weighted by the factors (weights) assigned to the search types according to the distribution of the type of device sought:

- Probability(a fuzzy search)= 0.9*weight(fuzzy)
- Probability(a precise search)=0.1*weight(precise)

The prediction system thus calculates a global score (step **17** in Figure 2) that indicates the type of search to be carried out in response to the received search query.

A search type context is obtained using:

- The search type probabilities deduced from previous search types sent by the initiator of the received search query, the previous search types being recovered from the search query history recovered using an identifier extracted from the received search query; and

- The distribution of the type of the device requested in the search query among the connected devices.

The search type prediction system weights the search type probabilities, deduced from the search history, using the distributions of connected devices located in the intelligent space. This makes it possible to weight the search history with more or less importance, according to the distribution of devices, for "precise" or "fuzzy" search algorithms.

This makes it possible, in particular, to determine whether a precise search algorithm or a statistical method (fuzzy search, leading to a less precise but more rapid result) should be used to respond to the search query.

Concerning the distribution of devices, four cases can be identified:

- The cardinality of the devices (that is to say the number of connected devices in the intelligent space 1) is low. Therefore, a precise search can be a priority (a search among a low number of devices is certainly rapid);
- The cardinality of devices is high, then to accelerate the search, a statistical approach may be adopted;
- The distribution of devices is Gaussian: some connected devices are of the same device category (that is to say an intelligent space comprising a plurality of lamps, or a plurality of telephones for example). A precise search may be adopted in this case to better distinguish among devices of the same category;
- The distribution of connected devices is stochastic: the objects are of highly varied types. In this case a fuzzy search may be adopted, the probability of distinguishing among devices of the same type thus being low.

Advantageously, the search query does not comprise search context parameters, in particular thresholds.

Predicting the search type to be carried out in searching for connected devices in the Internet of things, according to the method described above, presents several advantages. In particular, it makes it possible to,

- Adapt search results to a given context;
- Guarantee the best compromise between search precision and speed on the Internet of things;
- Consider the distribution of connected devices in the type of search to be carried out.

CLAIMS

1. A method for predicting the search type for a search query on the Internet of things comprising a step to estimate the distribution of the desired object using a means of calculation.

2. A method according to claim 1, characterised in that it also comprises a step to calculate search type probabilities based on the search types of previous search queries with the same identifier as said search query.

3. A method according to claim 2 characterised in that it also comprises a step to weight said probabilities with weights according to said distributions.

4. A method according to either of claims 2 or 3, characterised in that said previous search types are recovered from an internal database (11).

5. A method according to either of claims 2 or 3, characterised in that said previous search types are recovered from an intelligent sub-space near the initiator of said search query.

6. A search type prediction system in an intelligent space (1) comprising means adapted for estimating a distribution of the types of connected devices present in said intelligent space (1).

7. A system according to claim 6, characterised in that it also comprises a context analyser (10) adapted for extracting a query identifier and the type of device sought from a search query.

8. A system according to claim 7 characterised in that it also comprises the means to calculate search type probabilities from the search types of previous search queries having the same identifier as said search query.

9. A system according to claim 8 characterised in that it also comprises the means to weight said probabilities with weights according to said distributions.

10. A computer program product implemented on a memory medium, which may be implemented within a computer processing unit, and comprises instructions to implement the method according to claims 1 to 5.

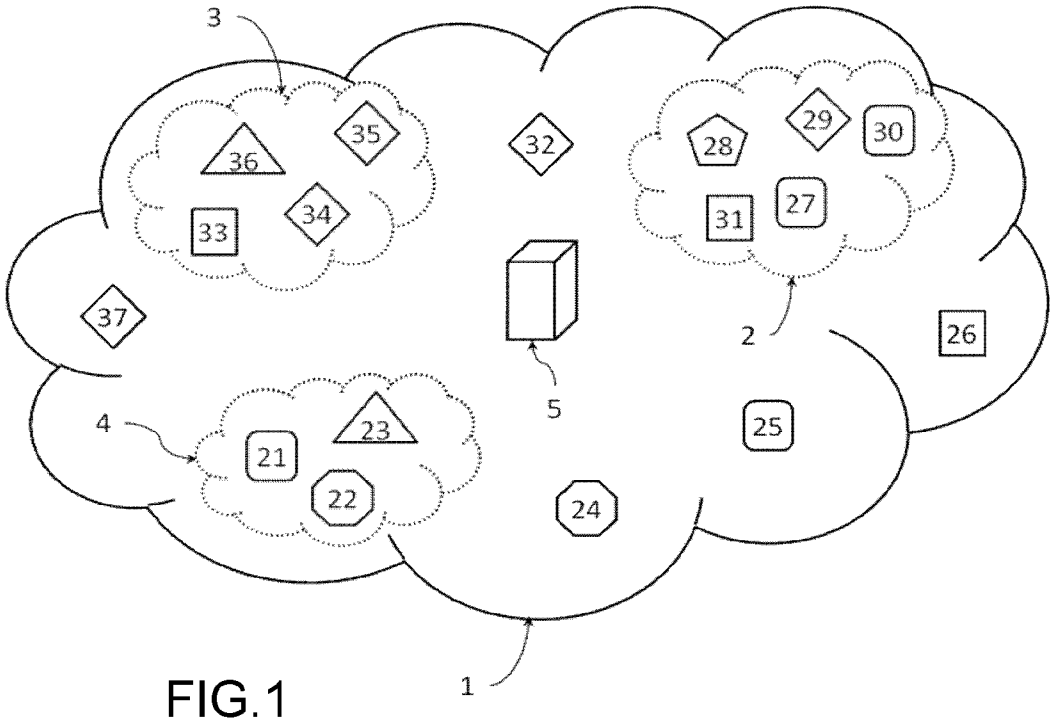
3
35
36
32
28
29
30
33
34
31
27
37
2
26
5
4
23
21
25
22
24
1

FIG.1

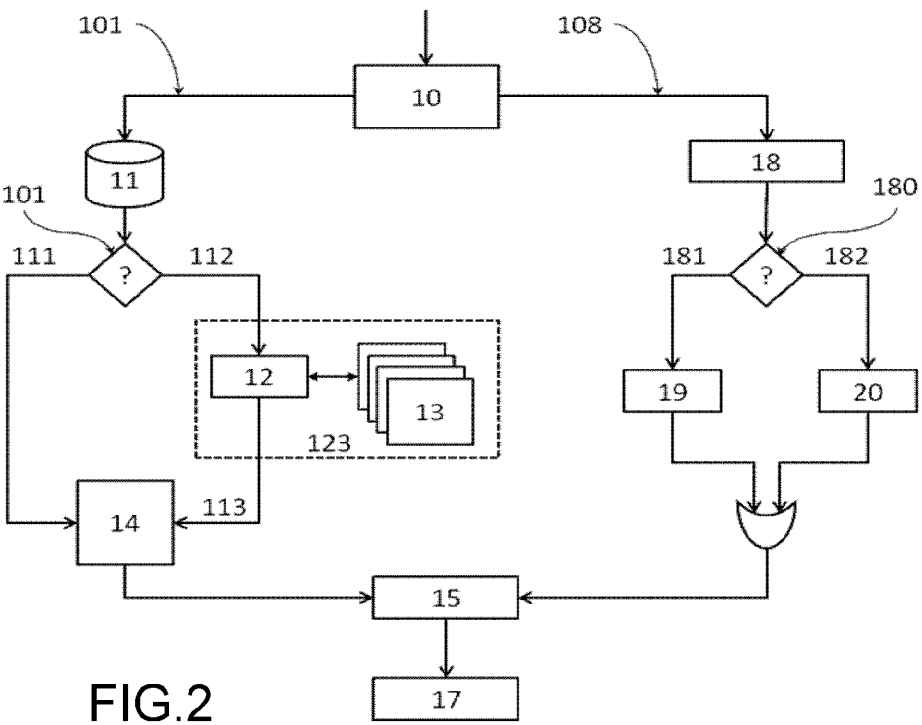
101
108
10
11
18
101
180
111
?
112
181
?
182
12
13
19
20
123
14
113
15
17

FIG.2

PATENT COOPERATION TREATY

PCT

DECLARATION OF NON-ESTABLISHMENT OF INTERNATIONAL SEARCH REPORT

(PCT Article 17(2)(a), Rules 13*ter*.1(c) and Rule 39)

Applicant's or agent's file reference	IMPORTANT DECLARATION	Date of mailing *(day/month/year)*
809628/JM		4 May 2012 (04-05-2012)
International application No.	International filing date *(day/month/year)*	(Earliest) Priority date *(day/month/year)*
PCT/EP2012/056086	3 April 2012 (03-04-2012)	6 May 2011 (06-05-2011)
International Patent Classification (IPC) or both national classification and IPC		
G06F17/30		
Applicant		
ALCATEL LUCENT		

This International Searching Authority hereby declares, according to Article 17(2)(a), that **no international search report will be established** on the international application for the reasons indicated below

1. ☐ The subject matter of the international application relates to:
   - a. ☐ scientific theories
   - b. ☐ mathematical theories
   - c. ☐ plant varieties
   - d. ☐ animal varieties
   - e. ☐ essentially biological processes for the production of plants and animals, other than microbiological processes and the products of such processes
   - f. ☐ schemes, rules or methods of doing business
   - g. ☐ schemes, rules or methods of performing purely mental acts
   - h. ☐ schemes, rules or methods of playing games
   - i. ☐ methods for treatment of the human body by surgery or therapy
   - j. ☐ methods for treatment of the animal body by surgery or therapy
   - k. ☐ diagnostic methods practised on the human or animal body
   - l. ☐ mere presentations of information
   - m. ☐ computer programs for which this International Searching Authority is not equipped to search prior art
2. ☒ The failure of the following parts of the international application to comply with prescribed requirements prevents a meaningful search from being carried out:
   ☒ the description ☒ the claims ☐ the drawings
3. ☐ A meaningful search could not be carried out without the sequence listing; the applicant did not, within the prescribed time limit:
   - ☐ furnish a sequence listing on paper complying with the standard provided for in Annex C of the Administrative Instructions, and such listing was not available to the International Searching Authority in a form and manner acceptable to it.
   - ☐ furnish a sequence listing in electronic form complying with the standard provided for in Annex C of the Administrative Instructions, and such listing was not available to the International Searching Authority in a form and manner acceptable to it.
   - ☐ pay the required late furnishing fee for the furnishing of a sequence listing in response to an invitation under Rule 13*ter*.1(a) or (b).
4. Further comments:

Name and mailing address of the International Searching Authority

European Patent Office, P.B. 5818 Patentlaan 2
NL-2280 HV Rijswijk
Tel. (+31-70) 340-2040
Fax: (+31-70) 340-3016

Authorized officer

GOLZE, Doreen
Tel: +49 (0)30 25901-704

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 203

The claims are not in compliance with the provisions of clarity of Art.6 PCT, as it is particularly burdensome for a skilled person to establish the subject-matter for which protection is sought.
The expression "types of connected devices" in claim 6 is too vague. What is meant? For example if the query mentions "Clio" is "Clio" the type of connected device or would it be "car"? In the case where the type has to be derived from the original query, how is this done?
Unfortunately there are no examples of query or types of connected devices given in the application.
The only passage of the description that could help to understand what exactly is meant with the "type" is p.7, l.15-18. Unfortunately this passage remains very vague and does not give any further details about what a "type" is exactly supposed to be.
The passage p.1 l.22 to p.2 l.7 does not describe enough which "entity" corresponds to a "type of connected device " as in claim 6. Indeed is "home appliance systems" a "type" of objects (and an object would then be e.g. an "alarm"), or is "alarms" a "type" of objects (and an object would be a particular instance "alarm A")?
Based on these passages it is not possible to define properly what is meant by a "type of object" or "type of device".
Concerning the document cited p.2 l.14-17 its content is not included in the application only because the document is cited. This cited document is therefore not part of the application. Moreover the use of the expressions "kind of products" or "kind of devices" used in the cited document does not solve the present clarity issues.
It is not clear what is meant by the "distribution of the types of connected devices" in claim 6. First this is due to the unclear meaning of the "types of connected devices" as explained above.
Moreover various expressions are used to describe this "distribution": the distribution of the desired object (p.3 l.15);the distribution of the type of device/object sought (p.7 l.16; p.8 l.5-6; p.8 l.20);the distribution of the types of connected devices (p.3 l.26; p.7 l.22-23; p.8 l.4; claim 6);the distributions of the type of device sought (p.8 l.14-15);the distribution of the type of the device requested in the search query among the connected devices (p.9 l.1-2);the distributions of connected devices (p.9 l.4-5);the distribution of devices (p.9 l.7,11 and 18);the distribution of connected devices (p.9 l.23; p.10 l.4).
It makes it unclear what is really meant: the distribution of the objects or the distribution of the types of objects? What is the link with the search query?
The only passage which could help to understand what is exactly the distribution is from p.7 l.18 to p.9 l.26
-p.7 l.18-20: this passage remains very vague and does not help to clearly define what is exactly the distribution.
-p.7 l.21-28: examples of distribution. This passage seems inconsistent: is this about the distribution of objects? of types of objects?
-p.8 l.4-13: what does a "test of the distribution of the types of connected devices" mean ? How is this test done? Two examples are provided (Gaussian distribution, stochastic distribution) but it is not clear what is the definition of the set on which the test is done. It is not clear either what the link between this set and the search query is.
-p.9 l.1-2: "the distribution of the type of the device requested in the search query among the connected devices" : the link is mentioned here between the distribution and the search query, however it is still not clear how this link is done.

-p.9, l.11-26: this passage gives a little bit more details about the possible distributions. However it is still not clear what the type of object is, therefore it is still not clear what is meant by the distribution. It is only written that a search (precise or fuzzy) "may be" adopted. The decision rules how to choose the search type (fuzzy/precise) based on the distribution remain undisclosed.
In conclusion there are 3 main clarity issues:
-there is no clear definition and no example of what a "type" of object or device is.
-there is no clear definition of what the distribution is.
-the link between the distribution and the search query is not clear. As a consequence there are doubts that choosing ("predicting") the search type (being fuzzy or precise) based on the distribution is really pertinent.
The invention is not disclosed in a manner sufficiently clear and complete to be carried out by a person skilled in the art, contrary to Art.5 PCT. The 10 pages are only a list of things that are supposed to be done, with no sufficient explanation of how they are done and interact with each other. There is no concrete example. It has been tried based on fig.2 to understand how the invention is supposed to work, with no success.
The non-compliance with the substantive provisions is to such an extent that a meaningful search of the whole claimed subject-matter could not be carried out (Art.17(2) PCT and Guidelines 9.30). There being no reasonable basis in the application that clearly indicates the subject-matter which might be expected to form the subject of the claims later in the procedure, no search at all was deemed possible.

The applicant's attention is drawn to the fact that claims relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure. If the application proceeds into the regional phase before the EPO, the applicant is reminded that a search may be carried out during examination before the EPO (see EPO Guideline C-VI, 8.2), should the problems which led to the Article 17(2) declaration be overcome.