Schreier G, Hayn D, Hörbst A, Ammenwerth E, editors. Proceedings of the eHealth2012. 2012 Mai 10-11; Vienna, Austria. OCG; 2012.

A REVIEW OF PUBMED ARTICLES RELATED TO MHEALTH USING TOPIC MODELLING

Kreiner K[1], Modre-Osprian R[1], Schreier G[1]

Abstract

Mobile applications for health are becoming more and more popular. However, there are still concerns, that mobile phone usage has adverse effects on health. We present a semi-supervised literature review of 1854 PubMed articles dealing with mobile phone usage and health. We hypothesize, that recent developments in mHealth resulted in a growing number of articles investigating beneficial effects of mobile phone usage on health. While the majority of articles still deals with adverse effects, articles reporting on beneficial effects show a significantly higher growth. Furthermore, we explore trending research topics and validate our review against a manual annotation done in 2007.

Keywords – ***mHealth, literature survey, mobile phones, topic modelling***

1. Introduction

"Mobile communication devices, in conjunction with Internet and social media, present opportunities to enhance disease prevention and management by extending health interventions beyond the reach of traditional care—an approach referred to as mHealth" [1] Especially the availability of cheap *smart phones* being equipped with a wide range of sensors as well as a convenient user interface has led to a growing interest of adopting this technology in numerous fields such as telemonitoring, personal health records or health diaries. In 2008 Apple and Google introduced *App Stores* where smart phone users can buy applications for their mobile phones. Since then, an increasing number of health applications is available. Today, Apple's App Store provides 9.000 health applications with the majority being either applications in the field of cardio fitness or diet. [2]

However, since years a heated debate is taking place in both media and the scientific community discussing adverse effects of mobile phone usage on health. Adverse effects range from exposure to radiation, interference with pacemakers, safety hazards while driving a vehicle or social implications of mobile phone usage.

In 2007 Schreier manually [3] reviewed 814 articles retrieved from the online platform PubMed[2]. Articles were flagged as *adverse, beneficial, irrelevant* or *duplicate.* Schreier concluded that 4 out of 5 articles discussed adverse effects of mobile phones on health. Adverse effects have been categorized into *exposure, driving, electromagnetic interference, social effects* or *other.*

[1] Safety and Security Department, AIT Austrian Institute of Technology GmbH, Graz

[2] http://www.ncbi.nlm.nih.gov/pubmed/

In this article we discuss, how recent developments in mHealth reflected in scientific literature related to this initial study. We therefore revisit the literature survey undertaken by [3], extend it to 2011 and hypothesize that the percentage of articles investigating beneficial effects of mobile phones in healthcare - as compared to all articles on mobile phones dealing with health issues - has increased over time.

However, given the exponentially increasing number of articles dealing with this subject over time, a manual method becomes more and more cumbersome. Therefore, we were looking for a way to automate the classification process.

2. Methods

2. 1. Data collection and pre-processing

We performed a keyword search for „(mobile [ti] OR cellular [ti]) AND (phone [ti] OR phones [ti] OR telephone [ti] OR telephones [ti])"on PubMed retrieving 1854 articles in November 2011. We then exported the results (consisting of title, abstract and meta-information) of the query as XML file.

In contrast to the review done by Schreier, we chose to create a topic model for the articles. Topic models are statistical models discovering abstract topics within a collection of documents based on distributions of words. For each document a distribution is created reflecting the probability that a given topic is found within the document.

We used MALLET [4], an open-source machine learning toolkit for language processing to create the topic model. In the first step we extracted titles and abstracts from the XML file as a base for the model.

MALLET requires knowing the number of topics in advance. We have experimented with 5, 10 and 20 topics. While using 5 and 10 topics led to poor results with topic boundaries being blurry, 20 topics proved to be a reasonable number of clearly defined topics.

MALLET creates a file showing the top keywords for each topic. *Table 1* shows two examples of top keywords:

Table 1: Keywords retrieved by the algorithm

Topic Number	Keywords
4	phone driving cell drivers hand performance free hands phones conversation
10	system patients patient care technology monitoring medical based management data diabetes

In the second step, a file was generated showing the distribution of topic probabilities for each article, i.e. PubMed abstract. We used this file and extracted the topic with the highest probability for a given topic as category for further investigations. We used this distribution to automatically assign each article to one topic. Furthermore we automatically grouped articles by year of publication.

2. 2. Finding topic titles

As [5] pointed out, interpretation of the topic top keywords and extracting meaningful titles (so-called topic model labels) is highly subjective and sometimes even difficult, since the top keywords may not be verbose at all. We chose a semi-supervised approach for creating labels.

First, we automatically created 2-word combinations (2-grams) of top-key words for each topic. We used those keywords to extract 3-word combinations (3-grams) as title candidates from the abstracts and counted their occurrences. For instance, the combination "ecg" and "phone" taken from the top keywords of a topic would produce "phone based ecg" as a title candidate.

We then automatically ranked the title candidates by number of occurrences. Based on these results we manually assigned an appropriate title for each topic from the top-10 list of title candidates. Each title was then classified as either "beneficial" or "adverse". *Figure 1* briefly summarises the process:

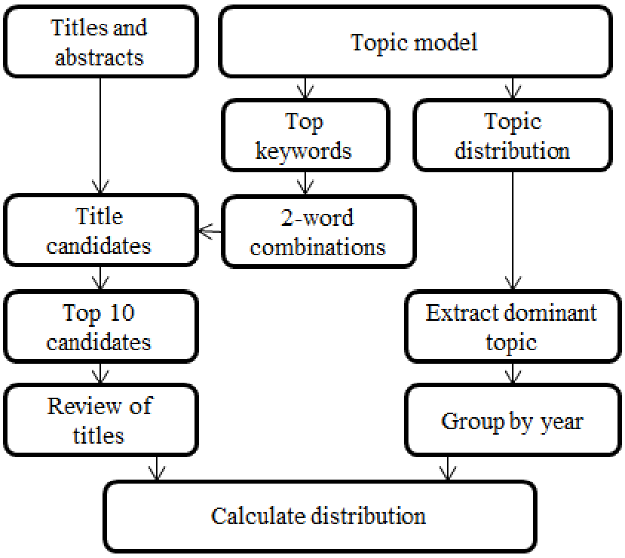

Figure 1: Creation and interpretation of the topic model. The review of titles has been done manually

3. Results

Table 2 illustrates the distribution of topics showing the total number of articles per topic and percentage in respect to the total number of articles:

3. 1. Validation

We validated our model against the manual annotation based on [3] which was manually extended to 1007 articles since publication. 877 articles were flagged either as *Beneficial* or Adverse. We compared those articles with topics assigned by the topic model reaching an agreement of 91% and a Cohen's kappa of 0.70 indicating *substantial agreement* according to [6].

Disagreements derive from the fact that the original annotation took into account only titles. However, some titles may be misleading while the abstracts itself clearly indicate adverse or beneficial effects. Further discrepancies stem from the fact that the topic model based classification, uses a simpler approach as compared to the human based classification. For example, it lacks a mechanism to identify Letters to the Editor in reaction to a preceding article and, as a consequence, counts this again as an article of the same category whereas the original annotation concept classified such articles as irrelevant ("duplicate").

Table 2: Topics found by the algorithm

Topic	Publications (until 2011)		Publications (until 2006)		Beneficial?
Mobile phones and brain tumours	155	8,30%	88	10,67%	
Electromagnetic field emissions	140	7,55%	67	8,13%	
Mobile phones and driving	128	6,90%	71	8,61%	
Pacemakers and electromagnetic interference	125	6,74%	95	11,52%	
Electromagnetic radiation of mobile phones	121	6,52%	58	7,03%	
Bacterial contamination of mobile phones	114	6,14%	66	8,00%	
Patient monitoring systems	111	5,98%	36	4,36%	Yes
Text messaging for health	111	5,98%	16	1,94%	Yes
Use of mobile phones in hospitals	97	5,23%	74	8,98%	
Cellular phone usage of Japanese high school students	94	5,07%	27	3,27%	
Exposure to radiation	87	4,69%	47	5,70%	
Mobile phones in developing countries	84	4,53%	34	4,12%	Yes
Health risks and mobile phone base stations	82	4,42%	37	4,49%	
Mobile phones and dermatitis	80	4,31%	27	3,27%	
Mobile phone cameras and health	73	3,93%	18	2,18%	Yes
Effects of radiation (animal experiments)	62	3,34%	15	1,82%	
Activity of daily living	58	3,12%	8	0,97%	Yes
Epidemiological studies and mobile phones	50	2,69%	17	2,06%	
Transmissions of ECGs	45	2,42%	11	1,33%	Yes
Patterns of mobile phone usage	37	1,19%	12	1,45%	
Total	**1854**		**824**		**482 (25,99%)**

Figure 2 illustrates the development of these topics over the course from 1990 to 2011. The number of publications dealing with adverse effects still outweighs the number of publications reporting beneficial effects. However, in 2011 the number of beneficial publications increased by 41% compared to 2010. In contrast the number of publications reporting adverse effects increased by 6,94%. Following figure illustrates the growth of articles reporting on beneficial or adverse effects since 1990:

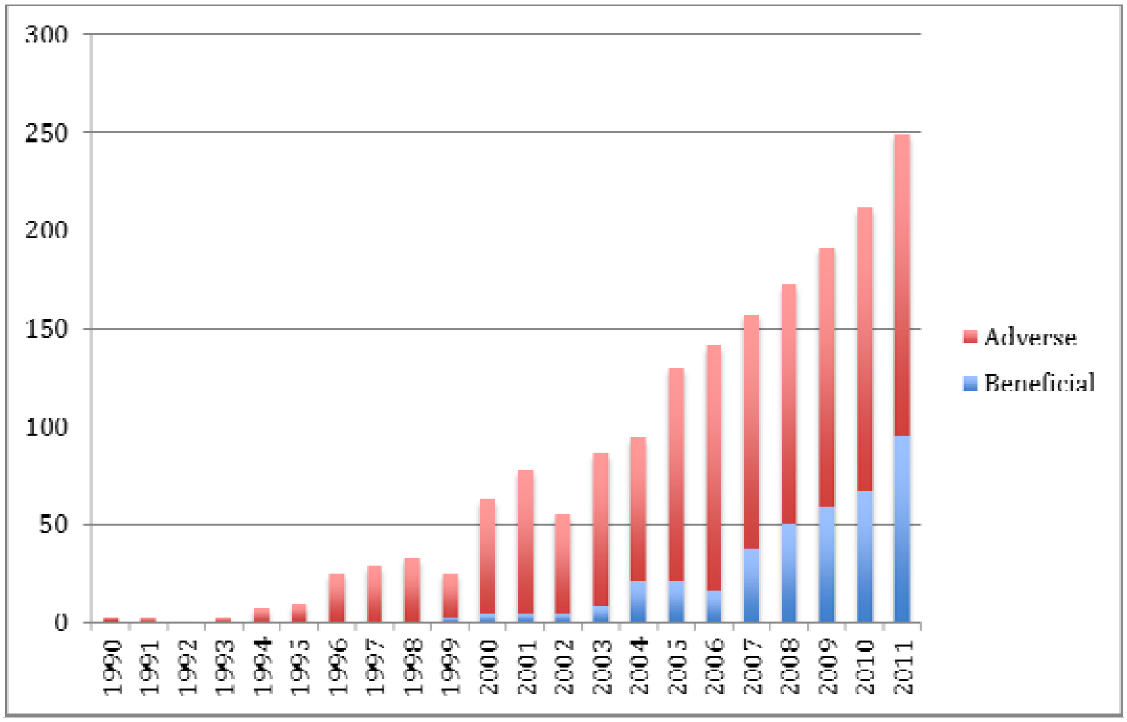

Figure 2: Number of beneficial and adverse effects of mobile phone usage since 1990

4. Discussion

Topics reporting adverse effects are generally growing slower relatively to the total number of articles compared to articles reporting beneficial effects. Especially articles on patient monitoring systems, short text messaging for health application and monitoring of activity of daily living are on the rise. However, articles reporting on radiation effects on animals still post high growth rates.

The topic model assigns a distribution of topics to each document. Therefore, topics are not mutually exclusive but overlapping. One could argue that using only the dominant topic blurs the results. This is especially true for documents that have a nearly equal distribution of topics.
Furthermore, it can be argued that a classification into *adverse* and *beneficial* effects is of little relevance. For instance, an article might discuss adverse effects of radiation but may conclude with insignificant results. Vice versa, a study might deal with telemonitoring as a beneficial aspect but conclude that a particular telemonitoring intervention does not have a significant positive impact on health outcome. In fact, the term *potential adverse effects* or *potential beneficial effects* respectively, might be more appropriate.

Finally, only one reviewer interpreted the topic model. It can be argued, that at least two reviewers should have interpreted key phrases and top keywords in order to extract meaningful titles. However, interpretation of most topics was straightforward, while two topics ("patterns of mobile use" and "Use of mobile phones in hospitals") proved to be particularly hard to interpret.

5. Conclusion and Outlook

We presented the development and validation of a semi-supervised literature survey of PubMed articles reporting on adverse or beneficial effects of mobile phone usage on health. While the majority of articles still deal with adverse effects, the number of articles reporting beneficial effects has dramatically grown within the last 5 years. This concept – which is basically not restricted to the present domain - may now be used to keep track with the development without the need to manually classify an exponentially increasing number of articles.

6. References

[1] Estrin D, Sim I, Health care delivery. Open mHealth architecture: an engine for health care innovation. Science. 2010 330:759-60

[2] Mobile Health News. Report: 13K iPhone consumer health apps in 2012 [homepage on the Internet]. [updated 2011 Sep 22; cited 2011 Jan 14]. Available from: http://mobihealthnews.com/13368/report-13k-iphone-consumer-health-apps-in-2012/

[3] Schreier G. The mobile phone and health – a Survey on Technologies, beneficial and adverse effects; Poster presented at Medinfo2007, Brisbane, Australia, 2007

[4] McCallum A.K. MALLET; A Machine learning for Language Toolkit [homepage on the Internet]. [updated 2011; cited 2011 Jan 22]. Available from: http://mallet.cs.umass.edu

[5] Mei, Qiaozhu and Shen, Xuehua and Zhai, ChengXiang. Automatic Labeling of multinominal topic models. In: Proceedings of the 13th ACM SIGKDD international conference on Knowledge Discovery and data mining.; San Jose, USA. ACM; 2007. p. 490-499

[6] Landis J R, Koch G G. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-174.

Corresponding Author
Karl Kreiner
Safety and Security Department, Austrian Institute of Technology GmbH
Donau-City Straße 1
A-1220 Vienna
Email: karl.kreiner@ait.ac.at