Delayed recognition

Sleeping Beauties ?

A JDIS publication !

Delayed recognition: recent developments and a proposal to study this phenomenon as a fuzzy concept

Ronald Rousseau

University of Antwerp, Faculty of Social Sciences, B-2020 Antwerp, Belgium

&

KU Leuven, Facultair Onderzoekscentrum ECOOM, Naamsestraat 61, Leuven B-3000, Belgium

ronald.rousseau@uantwerpen.be;
ronald.rousseau@kuleuven.be

Abstract

- Purpose – New developments in the study of delayed recognition are discussed.
- Design/methodology/approach – Based on these new developments a method is proposed to characterize delayed recognition as a fuzzy concept.
- Findings – A benchmark value of 0.333 corresponding with linear growth is obtained. Moreover, a case is discovered in which an expert found delayed recognition several years before citation analysis could discover this phenomenon.
- Research limitations – As all citation studies also this one is database dependent.
- Practical implications - Delayed recognition is turned into a fuzzy concept.
- Originality/value - The article presents a new way of studying delayed recognition.

Delayed recognition

- A publication suffering from delayed recognition is a publication that received very little attention shortly after publication, but received recognition later.

- If a publication goes through a period in which it receives no or almost no, citations it is often referred to as a Sleeping Beauty (van Raan).

Other names for this phenomenon

- Premature discovery,
- Suffering from Mendel's syndrome,
- Being a late bloomer
- Being ahead of one's time.

- But: Mendel's work was not totally unknown before the 20th century as was already mentioned by Garfield (1970).

Sleeping beauties?

- Kissing prince(s)
- Forming an ideal couple
- Male or female dominance
- Chastity (or not) of the sleeping beauty
- Beauty coefficient

- For this reason Sugimoto & Mostafa (2018) wrote an editorial, decrying this “clear violation of sociocultural norms”.

Possible solutions

- Late-Bloomer
- Hibernator - awakener (Rousseau)
- Sleeping paper

- Not “article-suffering-from-delayed-recognition”

The approach by Ke et al. (2015)

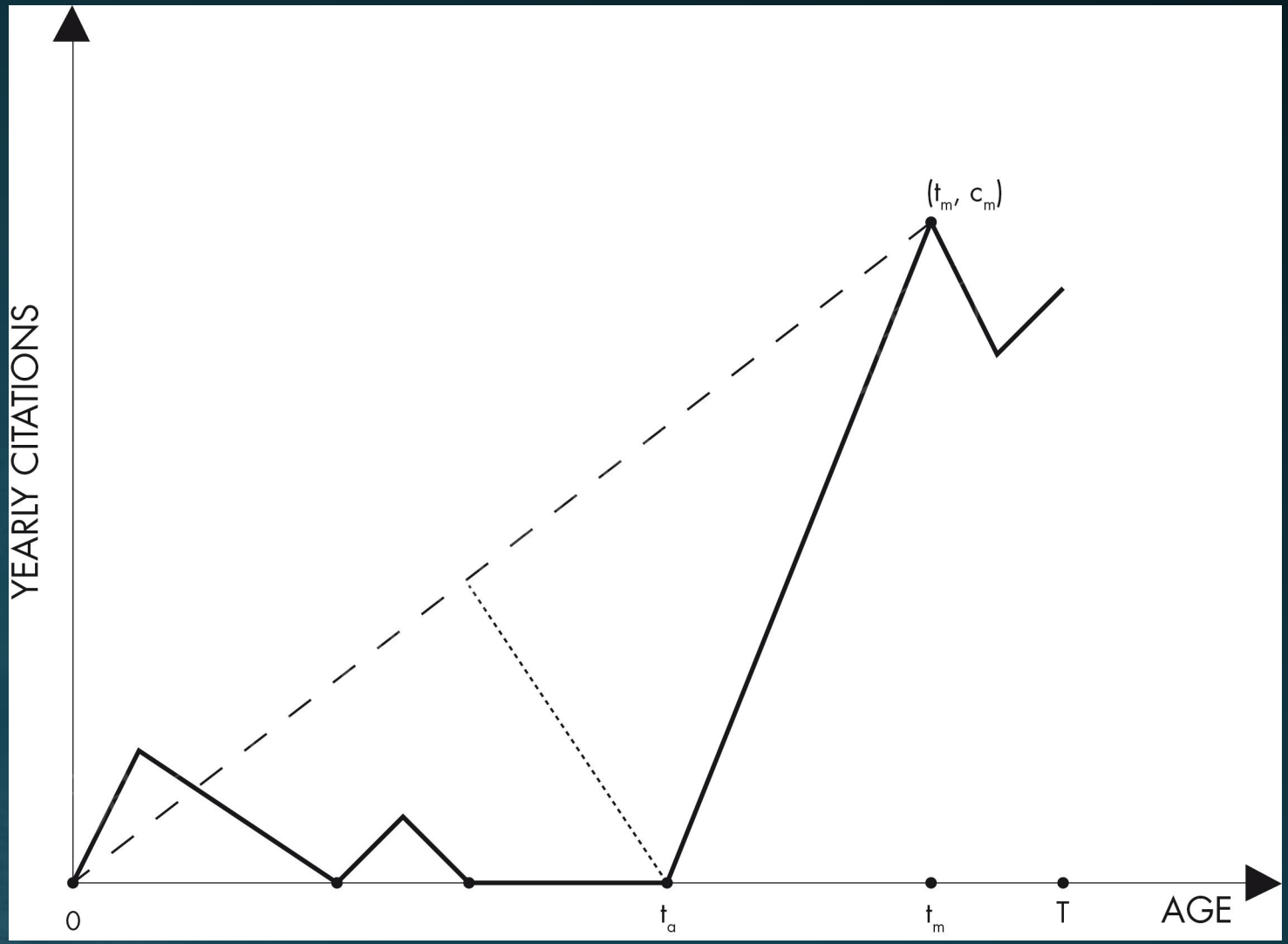
- Although being a sleeping beauty sounds like a yes/no situation, it is clear that delayed recognition is not a clear-cut phenomenon and a sleeping beauty in the eyes of one person may not be one in the eyes of a colleague. To solve this problem Ke et al. (2015) turned delayed recognition into a time-dependent continuous phenomenon by defining a beauty coefficient at time T, denoted as B(T).

$$y(t) = \frac{c_m - c(0)}{t_m} t + c(0)$$

Their method

- Let c(t) denote the yearly citation curve of an article, i.e., c(t) is the number of citations received in year t. The publication year is year t = 0 and t takes values between 0 and T.
- Let $c_m > 0$ be the maximum yearly number of received citations by this article, for which we assume that it happened in year t_m, with $0 < t_m \leq T$.
- The line y(t) connecting (0,c(0)) and the peak $(t_m,c_m) = (t_m,c(t_m))$, which is referred to as the recognition line, has equation:

$$y(t) = \frac{c_m - c(0)}{t_m} t + c(0)$$


YEARLY CITATIONS
(t_m, c_m)
0
t_a
t_m
T
AGE

The beauty coefficient B(T)

$$B(T)=\sum_{t=0}^{t_m}\left(\frac{\frac{c_m-c(0)}{t_m}t+c(0)-c(t)}{\max\{1,c(t)\}}\right)$$

- This is the normalized sum of the yearly differences between the citation curve and the recognition line.
- If the citation curve is concave, there is no delay and the result is negative.
- If the citation curve is convex the result is always positive
- If the citation curve is apprimatively linear the result is about zero.

The next step: Du Jian & Wu Yishan

- In recent papers Du & Wu (2017, 2018) note some disadvantages of the definition proposed by Ke et al. , the most important one being the high importance given to the peak. They claim that the determination of the B-value does not always work well. They, moreover, consider the role of the denominator in the original definition as just a way to avoid division by zero.
- For these reasons these authors propose a different approach, not based on the citation curve, c(t), but on the cumulative citation curve C(t).

My proposal: delayed recognition as a fuzzy concept

- I propose a framework to study delayed recognition of an article at a given moment in time, say T.
- More precisely, I consider the question: does this article suffers delayed recognition or has it in the past (while now it, perhaps, behaves like a normal article, already receiving a declining number of citations).
- Studying this question I consider three aspects: '**delayed**', '**recognition**' and **fuzzy membership**.

“Delayed”

- When it comes to the ‘delayed’ part, this implies that one must wait a certain period before one may say that there is a delay.
- In this study I wait at least ten years, but further investigations are needed to study the influence of this starting time.

“Recognition”

- Next I come to the ‘recognition’ part.
- I propose to concentrate on the 1% most cited publications in the same publication year as the publication under investigation.
- A choice must further be made to include all publication types in this 1% or only normal articles (or normal articles and reviews).
- I think that here all choices are valid, i.e., have some scientific value, but the choice must be stated clearly.

Fuzzy membership

- This value, between zero and one, must in a meaningful way express to which extent an article can be said to belong to the fuzzy set of publications with delayed recognition.
- This membership function, as calculated at time T, is denoted as DR(T).
- If an article is not ‘recognized’, i.e. it does not belong to the 1% most cited, it is not ahead of its time and its DR(T) value is set equal to 0 for any T. Our approach is based on ideas from Ke et al. and Du and Wu.

The framework

- I define a function K(t), with values between 0 and 1, for each year t and define DR(T) as shown below. The use of the Max function avoids that DR(T) decreases over time:

$$DR(T) = Max_{10 \le t \le T}\{0, K(t)\}$$

Construction of the function K(t)

I consider the line y(n) connecting the origin (0,0) with the point (t, C(t)). This line, which I call the recognition line at time t, has equation

$$y(n) = \frac{C(t)}{t}\, n$$

Now I calculate the sum of the differences in each n, 0 ≤ n ≤ t, between the line y(n) and the cumulative citation curve C(n). This sum is denoted as S(t).

$$S(t) = \sum_{n=0}^{t} \left(\frac{C(t)}{t}\, n - C(n) \right)$$

And finally:

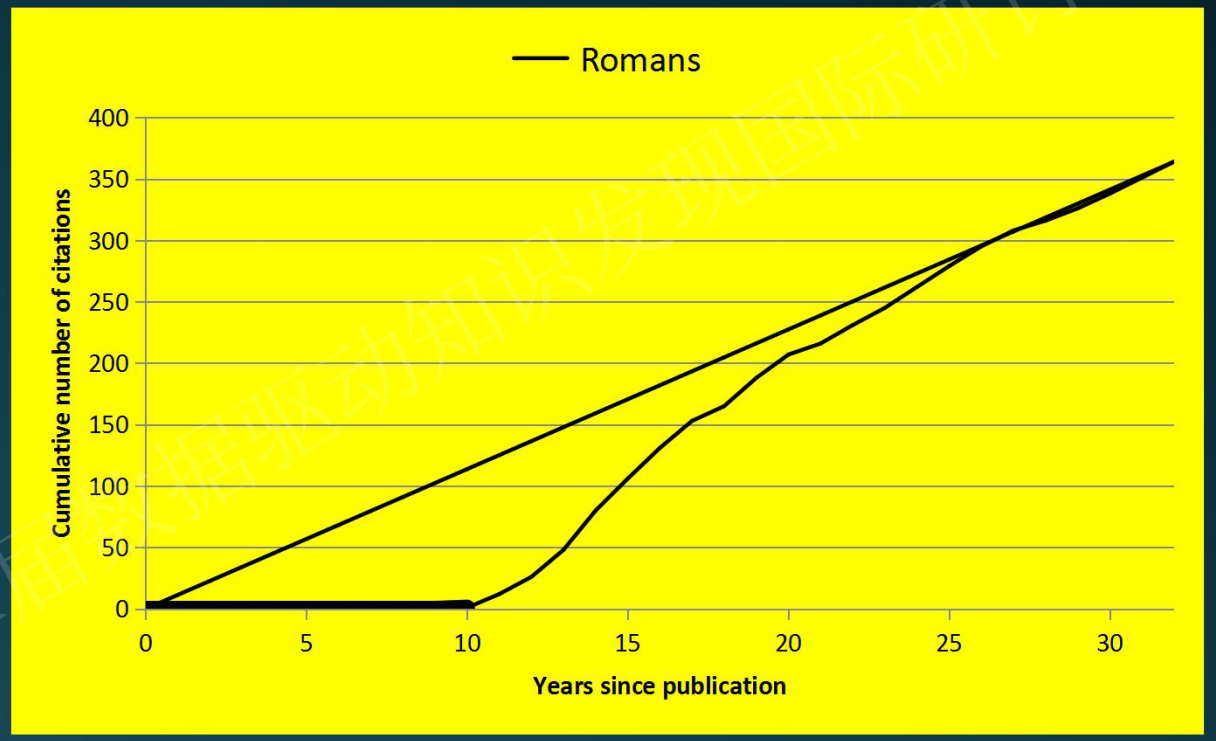
- K(t) is defined as the function S(t) divided by the maximum value at time t.
- This maximum value occurs when the publication receives its first citation in the year t.
- Because of this normalization K(t) has always values between –1 and +1.
- Recall now that

$$DR(T) = Max_{10 \le t \le T}\{0, K(t)\}$$

Theoretical examples

- If the cumulative citation curve is everywhere concave then K(t) is always negative and DR(T) = 0 for every T.
- Similarly, if c(n) is constant: c(n) = a > 0, then C(n) = a*n and the recognition line has equation y(n) = a*n. Clearly, also here DR(T) = 0, agreeing with the fact that there is no delayed recognition.
- If citations grow linearly in time, then DR(T)= 1/3. This result is in an important way different from the one obtained by Ke et al. B. Their B-value is zero in this case, although citations grow with time, indicating a delay in recognition.
- This value for linear growth can be considered as a benchmark value.

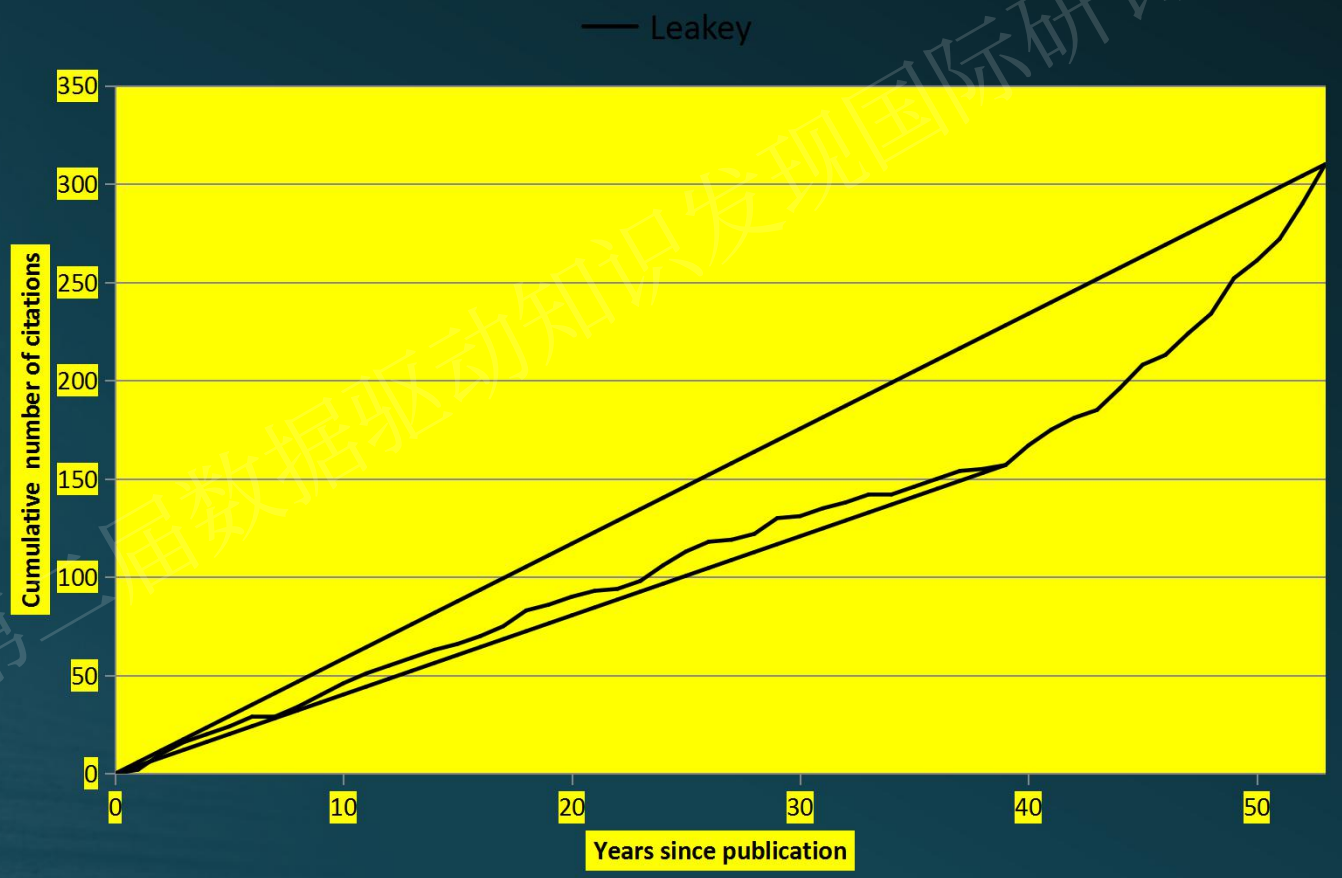
A real-world example: Romans (1986)



Comments and DR(T) value

- Its DR value is equal to 1.0 obtained in the year 1996, which is the first year for which we perform a calculation. Hence DR(T)= 1 for all T ≥ 10.

Another real-world example: Leakey et al. (1964)



Why did we study this article?

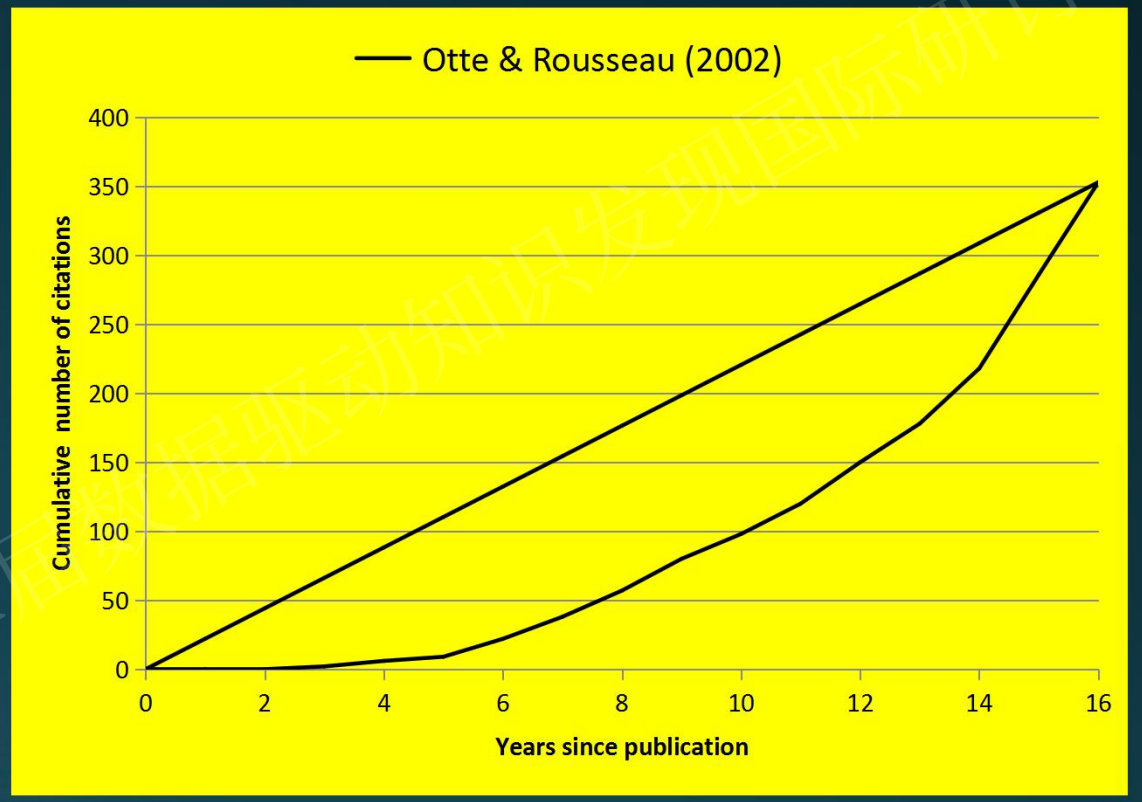
- Leakey, L.S.B., Tobias, P. V., & Napier, J. R. (1964). A new species of the genus Homo from Olduvai Gorge. Nature, 202(4927), 7-9.

- Tobias, P.V. (1996). Premature discoveries in science with especial reference to “Australopithecus” and “Homo Habilis”. Proceedings of the American Philosophical Society, 140(1), 49-64.

What did Tobias claim?

- In his paper from 1996 Tobias described how their findings were not accepted by their colleagues, but that step by step the original objections against their findings and corresponding theory fell away and, in his words, by 1984 their findings were accepted. This happened twenty years after their publication and hence, these findings were – rightly – described as a premature discovery.
- The citation curve does not show any sign of this observation. We think this illustrates the very important fact that using citations is just an operationalization (for delayed recognition) and experts may, rightly, have other opinions.

A last example: Otte & Rousseau (2002) on network theory; DR(2017) = 0.523



Conclusion

- I reviewed recent developments related to the study of delayed recognition, leading to the idea to consider delayed recognition as a fuzzy concept.
- I proposed a method to obtain fuzzy membership values. One of the requirements for suffering delayed recognition, is that the article must belong to the 1% most-cited ones. This means that at most 1% of the articles under consideration have a non-zero fuzzy membership value, and probably much less than 1%.
- The value 0.333 for linear growth in citations can be considered a benchmark for comparisons.

Conclusions (continued)

- Besides proper hibernators (sleeping beauties) who have a long period with no or few citations, articles suffering delayed recognition may have a convex cumulative citation curve, such as in the case of linear growth in citations. Examples of these two types are shown in this presentation: Romans (1986) being a proper hibernator and Leakey et al. (1964) and Otte & Rousseau (2002) being examples of the second type.
- We made the important observation that using citations to study delayed recognition is just a – convenient – operationalization of the concept, but that experts may agree on delayed recognition long before this is shown by citations. This is illustrated by the case of Leakey et al. (1964). This leads to the question: How good (adequate) is citation analysis for detecting premature discoveries?

Thank you for your attention

谢谢您的邀请。

谢谢您的莅临。