

# مقایسه تحلیل هم‌استنادی و تحلیل هم‌واژگانی در ترسیم نقشه کتابشناختی

## مطالعه موردی: دانشگاه تهران

نجمه سالمی\*  
دانشجوی دکتری علم اطلاعات و دانش‌شناسی  
دانشگاه تهران

کیوان کوشا  
دکتری علم اطلاعات و دانش‌شناسی؛ استادیار  
دانشگاه تهران

فصلنامه علمی پژوهشی  
پژوهشگاه علوم و فناوری اطلاعات ایران  
شاپا (جایی) ۸۲۲۲-۲۲۵۱  
شاپا (الکترونیکی) ۸۲۲۱-۲۲۵۱  
نمایه در SCOPUS, LISA و ISC  
<http://jimp.irandoc.ac.ir>  
دوره ۲۹ | شماره ۱ | صص ۲۵۳-۲۶۶  
پاییز ۱۳۹۲

دریافت: ۱۳۹۱/۰۶/۲۰ | پذیرش: ۱۳۹۱/۱۲/۱۹

**چکیده:** روش تحلیل هم‌استنادی و هم‌واژگانی، دو روش برای ترسیم نقشه‌های کتابشناختی هستند. نقشه‌های کتابشناختی نه تنها روند رشد حوزه‌های علمی را ترسیم می‌کنند بلکه زیربنای تصمیم‌گیری‌ها، برنامه‌ریزی‌ها و آینده‌نگری علمی به حساب می‌آیند. اگر محققان مختلف بر روی موضوعات پژوهشی و مفاهیم یکسانی تمرکز و مطالعه کنند، این انتظار به وجود می‌آید که در طیفی گسترده واژگان مشابهی را نیز برای مفاهیم مهم و یا مسائل حوزه تخصصی خود به کار برند؛ و همچنین همین تعداد پژوهشگر، هم‌زمان مجموعه مشخصی از مقالات پیشین خود را مورد استناد قرار دهند. از آنجا که مقالات حامل هم‌زمان واژه‌ها و مآخذ هستند، پژوهش حاضر سعی دارد دریابد آیا تکنیک بررسی مآخذ و بررسی واژگان در نتایج، مشابهت یا همگرایی نشان می‌دهد. جامعه تحقیق حاضر، کلیه مقالات دانشگاه تهران است که در بازه زمانی (۱۹۹۹-۲۰۰۸) در وب‌آوساینس نمایه شده‌اند. نتایج نشان داد که ۱۴ خوشه در نقشه هم‌استنادی و ۶ خوشه در نقشه هم‌واژگانی تشکیل شد. خوشه‌های حاصل از نظر اندازه بسیار متفاوت بودند. اغلب خوشه‌های نقشه هم‌استنادی در نقشه هم‌واژگانی در یک خوشه ادغام شد.

**کلیدواژه‌ها:** ترسیم نقشه کتابشناختی، تحلیل هم‌استنادی، تحلیل هم‌واژگانی، خوشه‌بندی

\* پدیدآور رابط:

najmehsalemi@yahoo.com  
I. kkoosha@ut.ac.ir

## ۱. مقدمه و بیان مسئله

نقشه‌های کتابشناختی نمایش بصری شبکه کتابشناختی هستند که در آن مجموعه‌ای از موجودیت‌های کتابشناختی و روابط میان آنها نمایش داده می‌شود و هدف آن ارائه‌ی نمایی از ساختار متون علمی در یک حوزه مورد نظر است (Van Eck and Waltman, 2010). نقشه‌های کتابشناختی می‌توانند حوزه‌های پژوهشی در یک رشته علمی را نشان دهند. همچنین با کمک این نقشه‌ها می‌توان میزان ارتباط حوزه‌های پژوهشی را نیز تعیین کرد. این نقشه‌ها اغلب زمانی که درک روابط میان مجموعه‌ای از داده‌ها مورد نظر است بسیار کارآمد هستند. نقشه‌های کتابشناختی ابزارهای مفیدی برای سیاستگذاری علم هستند. به این ترتیب که مطالعه موجودیت‌ها به تعیین شاخص‌ها کمک می‌کند و شاخص‌ها مبنای تصمیم‌گیری سیاستگذاران قرار می‌گیرند (Noyons 2001; Franklin and Johnston 1988; Healey and et. al 1986).

مرور پژوهش‌های پیشین نشان می‌دهد اسمال اولین کسی است که به ترسیم خوشه‌های مدارک پرداخت (Small and Griffith 1974; Small and Sweeney 1985). قدم اساسی در ساخت نقشه‌های حوزه‌ای را اسمال<sup>۱</sup> و مارشاکووا<sup>۲</sup> برداشتند. هر کدام از آنها هم‌استنادی میان مدارک را متغیر مناسبی برای نقشه علمی قلمداد کردند. اسمال معتقد بود که از طریق دسته‌بندی مدارکی که بسیار با هم استناد شده‌اند می‌توان ساختار علم را مطالعه کرد (Small 1973; Marshakova 1973). حرکت اسمال و مارشاکووا شروع ساخت نقشه‌های علمی شد که اغلب از نمایه استنادی علوم به‌عنوان منبع اطلاعات و هم‌استنادی استفاده می‌کردند. محققان دانشگاه در کسل (White and Griffith, 1981) نیز پیشگام نقشه‌های مؤلفان و مجلات بودند. فرانسویان (Callon and et. al 1983) نیز پیشگامان ترسیم نقشه‌های واژگان بودند که بعدها توسط پیترز و ون‌ران نیز پیگیری شد (Peters and Raan, 1993). در دهه ۹۰ روش‌های جدیدی در بازیابی اطلاعات و تکنیک‌های نو در تجزیه و تحلیل اطلاعات، ترسیم و حالت فضایی<sup>۳</sup> اطلاعات براساس تکنیک‌هایی برای ترسیم ساختار حوزه‌های علمی مورد مطالعه قرار گرفت (Moya-Anegon and et. al, 2004). نکته

---

1. Small  
2. Marshakova  
3. Spatial Positioning

مهم در مورد روش شناسی همه پژوهشگران این بود که گروهی از مدارک که زمینه موضوعی مشترکی داشتند مشخص می‌شدند. این امر گواه و شاهدهی بود بر آنکه علم شبکه‌ای است از حوزه‌های خاصی که با هم ارتباط درونی داشته و از طریق تحلیل‌های کمی تولیدات مکتوب می‌توان به آن ارتباطات درونی دست یافت. علم ماهیتی پویا دارد، جریان مداوم تولیدات علمی تغییرات مداومی نیز در ساختار علم ایجاد می‌کند. نقشه‌های کتابشناختی روند رشد و تغییرات حوزه‌های علمی را ترسیم می‌کنند.

به‌طور معمول مدارک، مؤلفان، مؤسسات، یا واژه‌ها واحد تحلیل در نقشه‌های کتابشناختی هستند. برای ترسیم نقشه کتابشناختی باید به نوع روابط میان موجودیت‌ها دقت کرد. بنابراین باید مقیاسی وجود داشته باشد تا میزان روابط میان موجودیت‌ها را بسنجد. به بیان دیگر باید دید دو موجودیت تا چه اندازه با هم مرتبط هستند، درک این ارتباط نیازمند یک مقیاس است. ذکر این نکته ضروری است که مقیاس اندازه‌گیری ارتباط برای واحدهای تحلیل (موجودیت‌ها) ممکن است متفاوت باشد (Van Eck and Waltman, 2007). رابطه مجلات، مدارک و مؤلفان اغلب براساس ارتباط استنادی بررسی می‌شوند. استاد مستقیم، هم‌استنادی، و زوج‌های کتابشناختی سه مقیاس برای سنجش ارتباط استنادی میان موجودیت‌ها هستند (Small, 1999). روش دیگر برای سنجش ارتباط میان داده‌های کتابشناختی، روابط هم‌نویسندگی است (White and Griffith 1981; White and McCain 1998). این مقیاس می‌تواند ارتباط میان مؤلفان، دانشگاه‌ها و کشورها را تعیین کند. این ارتباط براساس تعداد تألیفات مشترک میان آنها سنجیده می‌شود. روش معمول دیگر برای سنجش ارتباط میان مدارک و مؤلفان بررسی میزان ارتباط میان واژگان آنهاست (کالون، کورتایل، لاولیل<sup>۱</sup>، ۱۹۹۱). به این ترتیب می‌توان ارتباط دو موجودیت را براساس تعداد هم‌رخدادی واژگان آنها سنجید.

انواع مختلفی از نقشه‌های کتابشناختی را می‌توان ترسیم کرد و ساختار یک حوزه علمی را با آن نشان داد. برخی از نقشه‌ها بر پایه روابط هم‌استنادی و برخی براساس هم‌رخدادی واژه‌ها شکل می‌گیرند. خوشه مدارک هم‌استناد نمایانگر دانش پایه، مفاهیم کلیدی، روش‌ها یا تجارب محققانی است که آن خوشه را ساخته‌اند. در تحلیل

1. L. Laville

هم‌واژگانی، هم‌رخدادی کلیدواژه‌ها در عنوان، چکیده یا متن مقالات بررسی می‌شود. هم‌رخدادی کلیدواژه‌ها میزان ارتباط شناختی میان یک مجموعه مدارک را نشان می‌دهد. تفاوت مهمی که میان هم‌واژگانی و هم‌استنادی وجود دارد این است که در یک دوره پژوهشی مشخص، تحلیل هم‌استنادی به منابع استناددهنده (مقاله استناددهنده، مؤلفان استناددهنده) و مآخذ استناد شده (مؤلف استناد شده، مدرک استناد شده) نیازمند است؛ اما تحلیل هم‌واژگانی فقط نیازمند مجموعه‌ای از مقالات مجلات در یک حوزه موضوعی خاص است. با مقایسه نقشه‌های حاصل در دوره‌های زمانی مختلف، پویایی علم‌رديابی می‌شود (He 1999).

تحلیل هم‌استنادی و هم‌واژگانی دو روش معمول برای ساخت نقشه‌های استراتژیک و موضوعی یک حوزه‌اند. اگرچه روش تحلیل هم‌استنادی و هم‌واژگانی به‌طور جداگانه بسیار مورد توجه قرار گرفته‌اند، اما نباید فراموش کرد که هر دوی این روش‌ها - همچون تمامی روش‌های کمی دارای مسائلی هستند که نتایج را دچار تردید می‌کند. سؤالی که اینجا مطرح می‌شود این است که کدام روش در ارائه تصویر حوزه‌های علمی کارآمدتر است. به‌نظر می‌رسد به‌کارگیری هم‌زمان این دو روش می‌تواند به مقایسه الگوهای هم‌استنادی و هم‌واژگانی پرداخته و با جبران کمبودهای دو روش منظر تازه‌ای را در تحقیقات علم‌سنجی ایجاد کند. به این منظور در این تحقیق از روش تحلیل هم‌استنادی و تحلیل هم‌واژگانی برای ترسیم نقشه کتابشناختی استفاده شده است و نتایج حاصل به بحث گذاشته شده است.

## ۲. پرسش‌های پژوهش

۱. آیا خوشه‌های هم‌استنادی و خوشه‌های هم‌واژگانی تشکیل شده در شبکه علمی دانشگاه تهران همگونی دارند؟
۲. آیا تحلیل هم‌استنادی و هم‌واژگانی قابلیت بازنمایی حوزه‌های علمی را دارند؟

## ۳. جامعه تحقیق

در رتبه‌بندی‌های بین‌المللی از جمله نظام رتبه‌بندی شانگهای و سایماگو، دانشگاه

تهران در چند سال اخیر دارای رتبه اول کشوری است<sup>۱</sup>، بنابراین می‌تواند جامعه مناسبی برای مطالعات کتاب‌سنجی به حساب آید. جست‌وجو برای جمع‌آوری داده‌ها در پایگاه وب‌آوساینس<sup>۲</sup> انجام گرفت. در قسمت جست‌وجوی پیشرفته در سیستم جست‌وجوی پایگاه استنادی وب‌آوساینس برچسب‌هایی وجود دارد که با استفاده از آنها می‌توان به نتایج جامع‌تری دست یافت. ابتدا با فرمول  $CU = \text{Iran}$  نتایج برای مؤسسات و دانشگاه‌ها مورد بررسی قرار گرفت. در بخش جست‌وجوی پیشرفته فرمول جست‌وجوی زیر وارد شد. جست‌وجوی AD برای نام دانشگاه تهران و PY برای سال انتشار:  $AD = \text{University of Tehran AND PY} = 1999-2008$

از میان مدارک بازیابی شده در زمان جست‌وجوی اطلاعات، ۴۷۳۲ عنوان مقاله وجود داشت که به‌عنوان عنصر اصلی تحلیل انتخاب شدند. با در نظر گرفتن آستانه استنادی، ۹۰۰ عنوان مقاله که حداقل یک مؤلف وابسته به دانشگاه تهران داشته و تا تاریخ جست‌وجوی محقق حداقل ۱۰ استناد دریافت کرده بودند در جامعه پژوهش قرار گرفتند. ۱۱۲۲۴ کلیدواژه در این ۹۰۰ مقاله مشخص شد. مجموع مآخذ موجود در فهرست منابع مقالات اعضای هیئت علمی دانشگاه تهران نیز ۲۹۴۵۵ عنوان است. استنادها و واژه‌های تولیدات علمی دانشگاه تهران واحدهای تحلیل در این تحقیق به‌شمار می‌آیند. داده‌ها در فایل‌های ۵۰۰ تایی در فرمت صفحه‌گسترده استخراج و در کامپیوتر شخصی ذخیره شد سپس برای تحلیل وارد نرم‌افزار شد.

#### ۴. روش پژوهش

در این تحقیق از دو روش تحلیل هم‌استنادی و تحلیل هم‌واژگانی (هم‌رخدادی واژه) که از فنون اساسی در مطالعات علم‌سنجی است برای ترسیم نقشه علمی دانشگاه تهران در بازه زمانی ۱۹۹۹ - ۲۰۰۸ استفاده شده است و در نهایت نتایج حاصل از دو روش با هم مقایسه شدند.

برای تحلیل شبکه هم‌استنادی دانشگاه تهران از نرم‌افزار CiteSpace استفاده شد.

- 
1. QS World University Rankings (2010) at: <http://www.topuniversities.com/search/universities/university%20of%20Tehran?filters=type:profile>  
SIR World Report 2010: Regional Supplements at: <http://www.scimagoir.com/>
  2. Web Of Science

Citespace نرم‌افزاری تحت جاوا است که قابلیت مصورسازی و تحلیل روندهای موجود در متون را از طریق هم‌استنادی دارد. در مرحله اول خوشه‌های هم‌استنادی مقالات موضوعی تشکیل شد. در این تحقیق سعی شده است تا در تفکیک شبکه هم‌استنادی، مقالات مشابه در یک گروه قرار گیرند و تا حد امکان از ورود مقاله‌هایی با مشابهت کم جلوگیری شود. به این منظور عمل جداسازی شبکه  $G$  به زیرمجموعه‌های  $\{G_k\}$  چنین تعریف می‌شود:  $G = \bigcup_{k=1}^K G_k$  اگر  $G_i \cap G_j = \emptyset$  و زطل باشد. عملگر برش برای

$$\text{cut}(A, B) = \sum_{i \in A, j \in B} W_{ij}$$

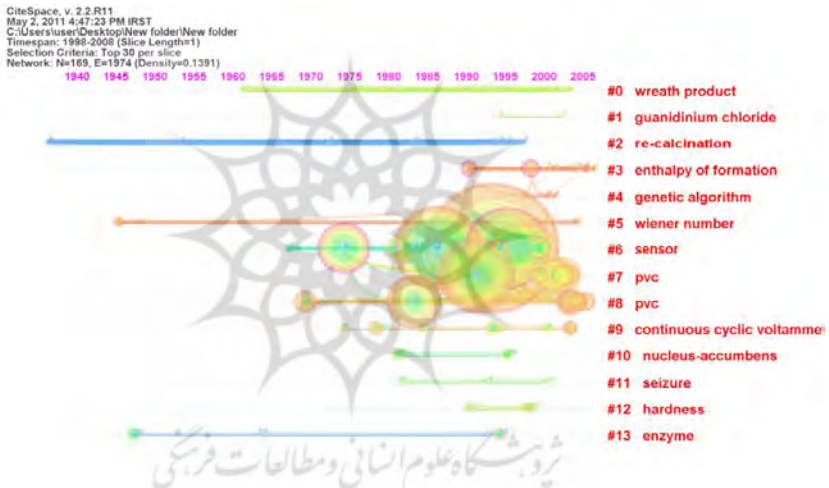
زیرمجموعه‌های  $A$  و  $B$  به این ترتیب است: سپس .  
برای خوشه‌های هم‌استنادی برچسب تعیین شد. دو الگوریتم رتبه‌بندی واژه *tf\*idf* (Salton and et.al 1975) و *log-likelihood ratio* (LLR) که در نرم‌افزار *citespace* طراحی شده، برچسب‌های خوشه‌ها را پیشنهاد می‌دهند. برچسب‌هایی که توسط الگوریتم *tf\*idf* انتخاب می‌شود معرف مفاهیم چشمگیر یک خوشه است. این وزن یک مقیاس آماری است که بسامد یک واژه مدرک را نسبت به مجموعه‌ای از مدارک نشان می‌دهد. هرچه دفعات حضور یک واژه در مدرک افزایش یابد اما در کل مجموعه متعادل بماند، آن واژه ارزش بیشتری برای توصیف مدرک دارد. اگرچه این روش وزندهی نسبتاً قدیمی است، اما سادگی و صحت نتایج، آن را به نقطه شروع روش‌های جدید تبدیل ساخته است (Salton and Buckley 1988). آزمون LLR یا لگاریتم نرخ مشابهت، احتمال وقوع داده‌ها در میان دو مدل را بررسی می‌کند. برچسب‌های LLR جنبه‌های منحصر به فرد یک خوشه را بزرگنمایی می‌کند. برای نام‌گذاری خوشه‌های هم‌استنادی از واژه‌های مقالات استاددهنده (مقالات دانشگاه تهران) استفاده شده است.

برای بررسی هم‌رخدادی واژگان مقالات دانشگاه تهران از نرم‌افزار *VOS Viewer* استفاده شد. *VOS Viewer* نرم‌افزاری است که به صورت اختصاصی برای ساخت و مصورسازی نقشه‌های کتابشناختی طراحی شده است. در این بخش، سه مرحله فیلترسازی و نایک (2010) به کار گرفته شد. در مرحله اول، مدارک از یک فیلتر زبانشناختی عبور کردند تا عبارات اسمی مشخص شوند. در مرحله دوم برای تشخیص یگانگی، عبارات

اسمی بررسی می‌شوند تا واحدهای معنایی مشخص شوند. در مرحله سوم، تمایل رخداد واژه‌ها به سمت موضوعات اندازه‌گیری می‌شود. در نهایت واژه‌های برگرفته از عنوان، چکیده و کدهای نمایه‌ای مقالات در خوشه‌های موضوعی قرار گرفتند.

## ۵. تحلیل داده‌ها

خوشه‌بندی مآخذ مقالات دانشگاه تهران به روش تحلیل هم‌استنادی نشان می‌دهد که ۱۴ خوشه اصلی در این شبکه تشکیل می‌شود (شکل ۱). برای تعیین برچسب خوشه‌های هم‌استنادی (پیوست ۱)، از کلیدواژه‌های نمایه‌ای استفاده شد. این انتخاب به‌طور خودکار توسط نرم‌افزار CiteSpace انجام شد.



شکل ۱. نقشه هم‌استنادی مآخذ

همانطور که در شکل ۱ دیده می‌شود خوشه #۲ قبل از دهه ۱۹۴۰ شکل گرفته و قبل از ۲۰۰۰ متوقف شده است. اگرچه این خوشه بزرگ نیست، اما گستره زمانی طولانی نزدیک به ۶۰ سال را طی کرده است. خوشه #۵ و #۱۳ در حدود ۱۹۴۵ شکل گرفته‌اند. سیر تاریخی این سه خوشه نشان می‌دهد موضوعات این خوشه‌ها اولین حوزه‌های علمی هستند که در شبکه هم‌استنادی دانشگاه تهران شکل گرفته‌اند. خوشه‌های #۶، #۷، و #۸

از دهه ۱۹۶۵ شروع به شکل‌گیری کردند و نزدیک به سال ۲۰۰۰ به اوج استنادی خود رسیده‌اند. ارتباطات درونی این خوشه‌ها حاکی از آن است که این خوشه‌ها دارای ارتباط موضوعی هستند. خوشه‌های ۱، ۱۰ و ۱۲ گستره زمانی کوتاهی نسبت به سایر خوشه‌ها دارند و می‌توانند مبین ایجاد حوزه‌های پژوهشی جدید در دانشگاه تهران باشند. سه خوشه بزرگ شبکه هم‌استنادی دانشگاه تهران (خوشه‌های ۶، ۷ و ۸) با محوریت الکتروشیمی به موضوعات تخصصی‌تر سنسورها، الکترودها و غشای مولکولی پرداخته‌اند. بیوشیمی مولکولی (خوشه ۱ و ۱۳)، شیمی مواد (خوشه ۲، ۳ و ۱۲)، طیف‌سنجی (خوشه ۴)، شیمی - ریاضی (خوشه ۵ و ۵)، الکتروشیمی (طیف‌سنجی) و کاربرد آن در داروسازی (خوشه ۹)، سم‌شناسی (خوشه ۱۰ و ۱۱)، حوزه‌های موضوعی نسبتاً کوچکی در مقالات دانشگاه تهران هستند.



شکل ۲. نقشه هم‌واژگانی مقالات دانشگاه تهران

نقشه هم‌رخدادی واژگان در مقالات دانشگاه تهران ۶ خوشه موضوعی را مشخص می‌کند (شکل ۲). شماره خوشه‌ها با توجه به بزرگی و تعداد موجودیت‌های آن تعلق گرفته است و موقعیت قرارگیری خوشه‌ها را فاصله موضوعی میان خوشه‌ها تعیین می‌کند. براساس این خوشه‌ها می‌توان حوزه‌های الکتروشیمی، شیمی تجزیه، بیوفیزیک - بیوشیمی،



طیف‌سنجی، سلامت و سم‌شناسی (داروسازی) را مشخص کرد. یافته‌های نقشه هم‌واژگانی نشان می‌دهد که خوشه‌ها از هم مجزا نبوده و با هم ارتباط مستقیم یا غیرمستقیم دارند.

**سؤال اول پژوهش:** آیا خوشه‌های هم‌استنادی و خوشه‌های هم‌واژگانی تشکیل شده در شبکه علمی دانشگاه تهران همگونی دارند؟

همانطور که دیده شد تعداد خوشه‌های هم‌استنادی و خوشه‌های هم‌واژگانی تشکیل شده در شبکه علمی دانشگاه تهران با هم متفاوت است. تعداد خوشه‌های تشکیل شده براساس هم‌استنادی ۱۴ خوشه و خوشه‌های هم‌واژگانی ۶ خوشه هستند.

برای مقایسه خوشه‌های هم‌استنادی با خوشه‌های هم‌واژگانی، واژه‌های پیشنهادی برای برچسب خوشه‌های هم‌استنادی (پیوست ۱) با واژه‌های موجود در خوشه‌های هم‌واژگانی مطابقت داده شد. جست‌وجوی واژه‌های خوشه‌های هم‌استنادی در نقشه هم‌واژگانی نشان می‌دهد سه خوشه هم‌استنادی ۶ و ۷ و ۸ (سنسور و پی وی سی) با موضوع الکتروشمی در خوشه ۱ نقشه هم‌واژگانی قرار گرفته‌اند. فاصله مکانی نزدیک این واژه‌ها در نقشه نیز دال بر همپوشانی موضوعی این مفاهیم است. این همپوشانی موضوعی در نقشه هم‌استنادی نیز با لینک‌های میان خوشه‌ها قابل تشخیص است. خوشه ۲ در نقشه هم‌واژگانی برچسب‌های خوشه‌های ۲، ۳ و ۱۲ نقشه هم‌استنادی را پوشش می‌دهد. این خوشه‌ها در موضوع طیف‌سنجی و شیمی مواد هستند. خوشه ۳ نقشه هم‌واژگانی خوشه ۹ نقشه هم‌استنادی را در خود جای می‌دهد. این خوشه در موضوع طیف‌سنجی است. پروفایل واژگانی خوشه ۴ در نقشه هم‌واژگانی با هیچ‌کدام از خوشه‌های هم‌استنادی همگونی ندارد. واژه‌های این خوشه از دو مقاله با موضوع «دیابت و مقاومت نسبت به انسولین» و «سندرم‌های متابولیکی در زنان و مردان» گرفته شده است. بدیهی است که این مقالات در بررسی هم‌واژگانی وارد شده‌اند، زیرا برخی از واژه‌های مقالات آستانه هم‌رخدادی را به دست آورده بودند، اما داده‌های لازم (مآخذ مشترک) برای خوشه‌بندی هم‌استنادی را نداشته‌اند.

خوشه ۵ نقشه هم‌واژگانی خوشه‌های ۹، ۱۰، ۱۱ و ۱۱ نقشه هم‌استنادی را پوشش می‌دهد. موضوع خوشه‌های ۱۰ و ۱۱ سم‌شناسی و خوشه ۹ الکتروشمی است. خوشه ۶ نقشه هم‌واژگانی معرف حوزه بیوشیمی مولکولی و همگون با خوشه‌های ۱ و ۱۳ هم

استنادی است.

**سؤال دوم پژوهش:** آیا تحلیل هم‌استنادی و هم‌واژگانی قابلیت بازنمایی حوزه‌های علمی را دارند؟

از آنجا که در تحلیل هم‌واژگانی آستانه هم‌رخدادی برای کلیدواژه‌ها تعیین می‌شود، حوزه‌های کوچک که واژه‌های آنها آستانه هم‌رخدادی را کسب نکرده بودند از نتایج حذف می‌شوند. برای مثال کلیدواژه‌های خوشه ۲# هم‌استنادی که ۴ مقاله دارد در نقشه موضوعی هم‌واژگانی دیده نشد. این مشکل در تحلیل هم‌واژگانی با پایین آوردن آستانه هم‌رخدادی کمتر می‌شود و کلیدواژه‌های بیشتری بازیابی می‌شوند، ولی حجم واژه‌های کم‌ارزش در نقشه به شدت بالا رفته و درک نقشه را مشکل می‌سازد.

از سوی دیگر در تحلیل هم‌واژگانی، تمام کلیدواژه‌های هم‌موضوع در یک خوشه قرار گرفته و تعداد خوشه‌ها نسبت به نقشه هم‌استنادی کمتر است. در مجموع باید گفت نقشه موضوعی اغلب حوزه‌ها را به شکل کلی‌تر ارائه داده است. نکته قابل توجه دیگر وجود خوشه کوچکی (خوشه ۴) در نقشه هم‌واژگانی است که کلیدواژه‌های آن با برچسب‌های پیشنهادی هیچ خوشه هم‌استنادی مطابقت نداشت. این امر می‌بین آن است که نقشه هم‌واژگانی توانسته این حوزه موضوعی را بازنمایی کند، اما در تحلیل هم‌استنادی این خوشه نادیده گرفته شده است.

## ۶. نتیجه‌گیری

شبکه ارتباطی در علم آنچنان گسترده و در حال تغییر است که مطالعه آن کاری دشوار است. برای رسیدن به درکی نسبی از این شبکه راهکارها و ابزارهای متعددی ایجاد شده که هر کدام جنبه خاصی از شبکه علمی را بازنمایی می‌کنند. هرگاه این راهکارها با هم ترکیب شود می‌تواند تصویری چندبُعدی از ساختار اجتماعی، دانش پایه و موضوعات پژوهشی ارائه دهد (Morris and Martens 2009). بررسی واژگان مدارک و استنادات، سال‌ها به‌عنوان ابزاری برای تعیین چارچوب حوزه‌های علمی به کار گرفته شده است. این تحقیق درصدد بود تا با اعمال هم‌زمان آنها قابلیت این دو روش را برای تعیین چارچوب حوزه‌های علمی مورد آزمون قرار دهد.

ترسیم نقشه شبکه علمی دانشگاه تهران نشان داد در فرآیند تشکیل خوشه‌های

هم‌استنادی در شبکه هم‌استنادی دانشگاه تهران هفت خوشه به صورت کاملاً مجزا دیده می‌شوند و یال‌های آنها فقط درون خوشه‌ای است. امتیاز ترسیم نقشه شبکه علمی در این بود که خوشه‌های مجزا به خوبی خودنمایی کردند. در این خوشه‌ها یک مرز تحقیقی وجود دارد که در طول دوره زمانی استنادی حفظ شده است. از سوی دیگر نقشه ترسیم شده فاصله خوشه‌ها را در یک فضای مجازی به تصویر کشید که در روش‌های سنتی قابل تشخیص نبود. درحقیقت مکان‌شناسی خوشه‌ها با استفاده از این روش بسیار آسان‌تر شد. سالتون (Salton 1971) معتقد بود از مآخذ می‌توان کلمات کلیدی برای توصیف مقالات استخراج کرد. این تحقیق مسیر کشف موضوعی را برعکس پیموده است. از آنجا که در وب‌آوساینس دسترسی به عنوان و چکیده مآخذ وجود ندارد، مقالات استناددهنده بار اطلاعاتی بیشتری را با خود حمل می‌کنند. بنابراین اگر مآخذ می‌توانند معرف موضوع یک مقاله باشند، مقاله نیز می‌تواند محتوای مآخذ خود را بازنمایی کند. به این ترتیب تحقیق حاضر تلاش کرد تا با استفاده از کلمات مقالات تولید شده در دانشگاه تهران، خوشه‌های هم‌استنادی مآخذ را در شبکه علمی دانشگاه تهران معین کند.

شیمی حوزه علمی بسیار وسیعی است و هر روز در تعامل بیشتری با سایر علوم قرار می‌گیرد. علاوه بر آن شیمی ارتباط تنگاتنگی با اقتصاد، صنعت و پزشکی دارد. برخی از مفاهیم در شیمی وجود دارد که پایه علم است و در تمام دنیا نیز بر آن تمرکز هست. برخی حوزه‌های علم شیمی (که بسیار زیادند) با توجه به نیاز جوامع و کشورها بیشتر مورد توجه قرار می‌گیرند و موضوعات در حال گسترش محسوب می‌شوند. مقالات دانشگاه تهران نشان می‌دهد تمرکز سه قطب پردیس علوم، دانشکده فنی و مؤسسه بیوشیمی و بیوفیزیک دانشگاه تهران بر روی سنسور است. سنسورها در الکتروشیمی بیوشیمی، الکتروشیمی بیوفیزیک، الکتروشیمی داروسازی، الکتروشیمی صنعتی و... تأثیرگذارند. مرکز عالی الکتروشیمی دانشگاه تهران نیز در زمینه سنسور برای علوم پایه و علم پزشکی فعالیت می‌کند. اگرچه شیمی در تولیدات دانشگاه تهران حرف اول را می‌زند، اما حوزه‌های مشخص شده در نقشه فقط بخش کوچکی از پایه علم شیمی را منعکس می‌کند و در حقیقت بیشتر خوشه‌های هم‌استنادی در نقشه هم‌واژگانی در هم آمیخته می‌شود. دلیل جدا شدن خوشه‌ها در نقشه هم‌استنادی، الگوی استناددهی آنهاست. همانطور که در شکل ۱ دیده می‌شود سه خوشه هم‌استنادی ۶، ۷ و ۸ هر سه در موضوع سنسور هستند

(pvc نوعی سنسور است). این سه خوشه در اثر خوداستنادی سه قطب علمی دانشگاه تهران ایجاد شده‌اند. بررسی مآخذ مقالات دانشگاه تهران نشان می‌دهد که یک حرکت چرخشی از تولید به مصرف در میان مقالات تولیدشده دانشگاه تهران وجود دارد. به این ترتیب مؤلفان دانشگاه تهران خود استناددهندگان بالقوه به تولیداتشان هستند. در این جمع کوچک مؤلفان، خوداستنادی فردی و گروهی چشمگیری قابل ردیابی است.

این الگوی استنادی از دو جنبه قابل بررسی است: اول اینکه یافته‌های دانشمندان ما در این حوزه‌های موضوعی کاملاً بومی است و می‌تواند پایه مطالعات بعدی قرار گیرد. فریم و نارین<sup>۱</sup> (Frame and Narin, 1977) معتقدند که منطقی است در یک کشور دانشمندانی را پیدا کنیم که نسبت زیادی از استنادهایشان را به هموطنان خود اختصاص دهند. این امر به دلیل میزان تلاش و برونداد علمی بالا در سطح بومی است. دوم، مؤلفان دانشگاه تهران استناد را معیار ارزشمندی قلمداد می‌کنند و تلاش دارند با خوداستنادی، نمایانی و ارزشمندی خود را بالا برند. از این منظر تحلیل هم‌استنادی می‌تواند بسیار آسیب‌پذیر باشد. به نظر می‌رسد تحقیقات بیشتری باید با ترکیب راهکارهای مختلف انجام شود تا خلأها یا توانایی‌های روش‌های ترسیم نقشه کتابشناختی آشکار شود.

دروندادهای مطالعات کتابسنجی اغلب در قالب رتبه‌بندی پژوهشگران و مؤسسات پژوهشی عرضه می‌شود. اما این سنج‌ها گویای کیفیت علمی نیست. اگرچه این رتبه‌بندی به سیاستگذاران کمک می‌کند تا نسبت به افزایش یا قطع حمایت از یک حوزه علمی یا مؤسسه تصمیم‌گیری کنند، اما در کل موجب ارتقاء عملکرد پژوهشی نمی‌شود. رتبه‌بندی به روابط بین موجودیت‌ها و چگونگی تأثیرگذاری این روابط بر تأثیر استنادی توجه ندارد. نتایج حاصل از این دو تحلیل اغلب به‌عنوان معیاری برای برنامه‌ریزی و سیاستگذاری در سطوح مدیریتی مورد استفاده قرار می‌گیرد. حذف یا افزایش اعتبارات پژوهشی، گسترش یا کاهش نیروی متخصص، تقویت حوزه‌های نوین و در حال رشد و... می‌تواند تحت تأثیر این داده‌ها قرار گیرد. بنابراین باید دید این دو روش تا چه حد می‌تواند در ارائه اطلاعات جامع به ما کمک کند. تحلیل هم‌استنادی و هم‌واژگانی همچون تمامی روش‌های کمی دارای مسائلی هستند که نتایج را دچار تردید می‌کند.

1. Frame & Narin

بنابراین به کارگیری همزمان دو روش قابل اعتمادتر است.

## ۷. پیشنهادات

پیشنهاد می‌شود تا تحقیق دیگری به بررسی ماهیت نقشه کتابشناختی دانشگاه‌های دیگر کشور پرداخته و میزان همگونی هم‌استادی در آن دانشگاه‌ها با دانشگاه تهران سنجیده شود. نتایج تحقیقات بعدی می‌تواند نشان دهد که آیا مقالات دانشگاه تهران نفوذی در مآخذ سایر محققان کشور دارند، و آیا مقالات دانشگاه تهران به چه میزان مورد استناد قرار می‌گیرند؟

پیشنهاد می‌شود تحقیق دیگری حضور اعضای هیئت علمی دانشگاه تهران را در پایگاه‌های دیگر همچون اسکوپوس یا مجلات دسترسی آزاد بررسی و با ترسیم نقشه کتابشناختی، نتایج را با این پژوهش مقایسه و ارزیابی کند.

## ۸. منابع

- Callon, Michel; Courtial, Jean Pierre; and Laville, Francisco. 1991. Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry. *Scientometrics* 22(1): 155-205.
- Callon, Michel; Courtial, Jean Pierre; Turner, William; and Bauin, Serge. 1983. From Translation to Problematic Networks: An Introduction to Co-Word Analysis. *Social Science Information*. 22(2):191-235.
- Frame, J. Davidson; Narin, Francis; Carpenter, Mark P. 1977. The distribution of world science. *Social studies of science*. 7(4): 501-516.
- Franklin, J. J., and Johnston, R. 1988. Co-citation bibliometric modeling as a tool for S&T and R&D management: Issues, applications, and developments. In Anthony, F. J. Van Raan (Ed.), *Handbook of quantitative science and technology research* (325-389). Elsevier Science Publishers.
- He, Qin. 1999 "Knowledge Discovery Through Co-Word Analysis". *Library Trends*. 48(1): 133-59
- Healey, Peter; Rothman, Harry; and Hoch, Paul. K. 1986. An experiment in science mapping for research planning. *Research Policy* 15(5): 233-251.
- Marshakova, Irena. 1973. System of document connection based on references. *Nauchno-Tekhnicheskaya Informatsiya*. 2(6): 3-8.
- Morris, Steven; and Martens, Betsy Van der Veer Martens. 2009. Mapping research specialties. *Annual Review of Information Science and Technology* 42(1): 213-295.
- Moya-Anegon, Felix de, and [et. al.]. 2104. A new technique for building maps of large scientific domains based on the cocitation of classes and categories. *Scientometricst* 61(1):129° 145.
- Noyons, Ed. 2001. Bibliometric mapping of science in a science policy context. *Scientometrics*. 50(1): 83-98

- Peters, H.P.F.; van Raan, Anthony.F.J. 1993. Co-word-based science maps of chemical engineering. Part II: Representations by direct multidimensional scaling . *Research Policy* 22(1): 23-45.
- Salton, Gerald; Buckley, Christopher. 1988. Term-weighting approaches in automatic text retrieval . *INFORMATION PROCESSING AND MANAGEMENT*.24(5): 513-553
- Salton, Gerard; Allan, James; Buckley, Chris; Singhal, Amit. 1975. Automatic analysis, theme generation and summarization of machine-readable texts . *Science* 264(5164): 1421-1426.
- Small, H. 1973. Co-citation in the scientific literature: a new measure of the relationship between two documents . *Journal of the American Society for Information Science and Technology*, 24(4): 265° 269.
- Small, Henry. 1999. Visualizing science by citation mapping. *Journal of the American Society for Information Science* 50(9): 799-813.
- Small, Henry., & Griffith, Belver. C. 1974. The structure of scientific literatures I: Identifying Bibliography 215 and graphing specialties. *Science Studies* 4(1): 17-40
- Small, Henry, & Sweeney, E. 1985. Clustering the science citation index using cocitations. I. A comparison of methods. *Scientometrics* 7(3° 6): 391-409.
- Van Eck, Nees Jan and Waltman, Ludo. 2007. VOS: A new method for visualizing similarities between objects. In H.-J. Lenz & R. Decker (Eds.), *Advances in data analysis: Proceedings of the 30th annual conference of the german classification society* (p. 299-306). Springer.
- Van Eck, Nees Jan, and Waltman, Ludo. 2010. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* 84(2): 523-538.
- White, Howard. D., and Griffith, Belver. C. 1981. Author cocitation: A literature measure of intellectual structure. *Journal of the American Society for Information Science* 32(3): 163-171.
- White, Howard D., and McCain, Katherine W. 1998. Visualizing a discipline: An author co-citation analysis of information science, 1972° 1995. *Journal of the American Society for Information Science* 49(4): 327-355.

پژوهشگاه علوم انسانی و مطالعات فرهنگی  
پرتال جامع علوم انسانی

# Co-citation Analysis and Co-word Analysis in Bibliometrics Mapping: A Methodological Evaluation

**Najmeh Salemi\***

PhD Student of Information Science, University of Tehran, Iran

**Keyvan Koosha<sup>1</sup>**

Former Assistant Professor of Information Science, University of Tehran, Iran

Iranian Journal of  
**Information  
Processing &  
Management**

**Abstract:** The general aim of a bibliometric map is to provide an overview of the structure of the scientific literature in a certain domain or on a certain topic. In the context of science policy and research management, bibliometric maps can be used to support decision making by governments, funding agencies, and universities. The articles have words and references which are the bibliometric data for co-citation analysis and co-word analysis. The two found bibliometric approaches whether they show any correlation? These approaches were applied to University of Tehran productions published between 1999 and 2008, retrieved from the Web of Science. The results reveal 14 co-citation clusters and 6 co-word clusters. The clusters vary considerably in size. Co-word clusters emerged co-citation clusters.

**Keywords:** Bibliometric mapping, Co-citation analysis, Co-word analysis, Clustering.

Iranian Research Institute Iranian  
for Science and Technology

ISSN 2251-8223

eISSN 2251-8231

Indexed in LISA, SCOPUS & ISC

Vol.29 | No.1 | pp: 253-266

Autumn 2013

\* Corresponding Author:  
najmehsalemi@yahoo.com  
1. kkoosha@ut.ac.ir