

خوشه‌بندی استان‌های ایران بر پایه

معیارهای شکاف دیجیتال به کمک روش K-MEANS

احمد یوسفان^۱، الهام یوسفیان^۲

^۱ مربی، گروه مهندسی کامپیوتر، دانشگاه کاشان، کاشان، ایران

yoosofan@kashanu.ac.ir

^۲ دانش‌آموخته مهندسی کامپیوتر، دانشگاه کاشان، کاشان، ایران

elhamusf@yahoo.com

چکیده: در این مقاله، مفهوم شکاف دیجیتال و تعدادی از روش‌های تحلیل شکاف دیجیتال توضیح داده شده است. روش‌های تحلیل شکاف دیجیتال، نمایه نامیده شده‌اند. نمایه‌های گوناگون، نشانگرهای گوناگون و فرمول‌های محاسباتی متفاوت دارند. گردآوری داده‌های نشانگرهای شکاف دیجیتال می‌تواند تا اندازه‌ای سخت باشد و این مسئله‌ای بنیادی برای محاسبه یک نمایه است. نگارندگان، برخی از نشانگرهای استان‌های ایران را گردآوری و محاسبه کرده‌اند، اما این نشانگرها برای محاسبه یک نمایه استاندارد کافی نبودند. این نشانگرها به صورت حادی، شکاف ژرف میان استان‌ها را نشان می‌دادند. برای نشان دادن دقیق‌تر نابرابری اجتماعی در به کارگیری فناوری اطلاعات و ارتباطات در میان استان‌های ایران، از الگوریتم خوشه‌بندی مشهور k-means بر روی نشانگرهای استان‌ها استفاده شده است. نتایج خوشه‌بندی به خوبی نشان داد که تهران، وضعیت یکتایی در میان دیگر استان‌های کشور دارد، زیرا تهران همواره به تنهایی در خوشه جداگانه‌ای گذاشته می‌شود. این نتیجه بدان معنی است که فناوری اطلاعات به شکل ناعادلانه‌ای میان استان‌های ایران پخش شده است.

واژه‌های کلیدی: شکاف دیجیتال، فناوری اطلاعات، ایران، آمادگی الکترونیکی، خوشه‌بندی.

۱. مقدمه

چندانی انجام نشده است. مهم‌ترین کاری که پیش از این انجام شده، [۸] است. البته در زمینه کلی تمرکزگرایی در ایران، کارهای زیادی مانند [۹] تاکنون انجام شده است. همچنین در زمینه شکاف دیجیتال و بررسی آمادگی الکترونیکی در بخش‌های گوناگون کارهایی انجام شده است، مانند مقایسه دبیرستان‌های استان تهران [۱۰]، ارزیابی آمادگی الکترونیکی دانشگاه‌ها [۱۱]، طبقه‌بندی پارک‌های فناوری [۱۲]، مدلی برای سنجش آمادگی الکترونیکی شرکت‌های کوچک و متوسط [۱۳ و ۱۴] و آمادگی الکترونیکی در منابع انسانی (شرکت نفت) [۱۵].

۳. شکاف دیجیتال

در مسیر توسعه اجتماعی با ورود فناوری اطلاعات و ارتباطات (ICT)، نیازهای بشری صرفاً در اقتصاد و معاش خلاصه نمی‌شود، بلکه هر روز نیازهای بشری فراتر از معیشت روزمره نمایان می‌گردد. امروزه انسان برای تحکیم آسایش و رفاه خود، نیازمند دانش و اطلاعات است. در بیشتر جامعه‌ها حق ایجاد اطلاعات به طور عملی در اختیار همگان قرار می‌گیرد، اما توزیع یا به اشتراک‌گذاری، از آنان گرفته می‌شود؛ از این رو، منابع اطلاعاتی به طور یکسان مورد بهره‌برداری جامعه قرار نمی‌گیرد. یک جامعه سالم در صورتی می‌تواند رفاه عادلانه برای افراد خود فراهم کند که افزون بر منابع اقتصادی، منابع اطلاعاتی را نیز به طور یکسان در اختیار همگان بگذارد [۱۶]. جامعه‌های غیر دموکراتیک از نظر فناوری اطلاعات، دچار شکاف دیجیتالی ژرفی هستند. شکاف دیجیتالی، فرآورده پخش ناعادلانه اطلاعات در پهنه اجتماع است و به موجب آن، اعضای جامعه برای تصمیم‌گیری و برنامه‌ریزی، به طور برابر به دانش و اطلاعات کارآمد دسترسی ندارند. در این صورت، افزون بر اینکه در منابع مادی با فقیر و غنی روبه‌رو هستیم، در منابع اطلاعاتی نیز با فقر اطلاعاتی^۲ و ثروت اطلاعاتی^۳ مواجهیم [۱۷ و ۱۸].

شکاف دیجیتال برای ایران را می‌توان از دو جنبه بررسی کرد: از یک سو، شکاف دیجیتال میان ایران و دیگر کشورها، که

به فناوری گردآوری، ذخیره‌سازی، ساختاربندی، مدیریت، فشرده‌سازی، انتقال اطلاعات و در نهایت، پردازش، دستیابی و تفسیر این اطلاعات، فناوری اطلاعات گفته می‌شود [۱]. امروزه فقط ۰/۰۰۳ از اطلاعاتی که در دنیا به صورت سالانه تولید می‌شود، به صورت چاپی می‌باشد [۲] و اطلاعات دیجیتالی به صورت چشمگیری در حال افزایش است [۳]. این شتاب در پیشرفت فناوری اطلاعات به شکل‌های گوناگونی بر زندگی انسان تأثیر گذاشته است به گونه‌ای که ارائه زنده پدیده‌ها، تصویر، فیلم، ویدئو و... به بخشی از زندگی روزمره ما بدل شده است [۴]. اکنون اطلاعات، یک کالای ارزنده اقتصادی شده است و مؤسسات اطلاع‌رسانی از فناوری پیشرفته‌ای در زمینه جمع‌آوری، انباشت، پردازش و اشاعه اطلاعات برخوردارند. در واقع، امروزه اطلاعات، یکی از با ارزش‌ترین منابع سازمان‌ها به حساب می‌آید. در دنیای امروزی، تصمیم‌گیری‌های صحیح، مؤثر، سازنده، دقیق و به موقع مدیران، به اطلاعات یکپارچه وابسته است [۵].

به گمان برخی از پژوهشگران، ایران در یک دهه گذشته، پیشرفت بسیار خوبی در زمینه فناوری اطلاعات داشته است [۶]. همچنین برخی بر این باورند که فناوری اطلاعات به دلیل نو بودن و شتاب بسیار زیاد آن، فرصتی برای کشورمان فراهم می‌کند تا بتوانیم بدون توجه به عقب‌ماندگی گذشته به پیشرفت بزرگی دست یابیم [۷].

در این مقاله، کوشش شده تا شکاف دیجیتالی^۱ به شکلی رسمی توضیح داده شود و برخی از روش‌های کمی بررسی آن معرفی گردد. هر کدام از این روش‌ها دربردارنده تعدادی نشانگر هستند که باید به صورت عددی گردآوری شوند. برخی از این داده‌های کمی را نمی‌توان در کشور گردآوری کرد و گردآوری برخی نیز بسیار دشوار است. در ادامه بر روی این داده‌ها، داده‌کاوی انجام شده تا به شکل بهتری بتوان نتیجه‌گیری کرد.

۲. گذری بر پژوهش‌های پیشین

در زمینه شکاف دیجیتال در میان استان‌های ایران، کار پژوهشی

2. Information Poverty
3. Information Wealth

1. Digital divide

برخی از آمارهای رسمی و خبرها نشان از افزایش آن دارد [۱۹ و ۲۰]؛ از سوی دیگر، شکاف دیجیتال میان استان‌های کشور، که در این مقاله به آن پرداخته شده است.

۴. روش‌های ارزیابی شکاف دیجیتال

امروزه مؤسسه‌های گوناگونی بر روی ارزیابی گسترش فناوری اطلاعات و شکاف دیجیتال مطالعه می‌کنند که هر کدام از آن‌ها، نشانگرها^۱ و دسته‌بندی ویژه خود را دارند که نمایه^۲ نامیده می‌شود. نشانگرها و دسته‌های مشترکی نیز میان نمایه‌های گوناگون وجود دارد. تفاوت میان نمایه‌ها بیشتر برگرفته از تعریف و نوع نگاهی است که به این مقوله دارند؛ برای نمونه، اتحادیه بین‌المللی مخابرات، ITU، در یکی از نمایه‌های خود بیشتر به نشانگرهای مخابراتی اهمیت داده است.

برخی از پژوهشگران نیز، بسیاری از روش‌های ارزیابی شکاف دیجیتال را کافی نمی‌دانند و گاه آن‌ها را نادرست می‌دانند؛ برای نمونه، در این روش‌ها به این موضوع که افراد چه کاری می‌توانند با رایانه انجام دهند، پرداخته نمی‌شود [۲۱]؛ به عبارتی، مهارت کاربر چندان در نظر گرفته نمی‌شود. دقت شود که اغلب، مدرک نشان‌دهنده میزان کار واقعی و سودمندی که یک فرد با رایانه انجام می‌دهد، نیست. مقاله [۲۲] یک بررسی تحلیلی بر روی چند نوع از این روش‌های ارزیابی انجام داده است. نمایه‌های گوناگون دیگری نیز وجود دارد یا پیشنهاد شده است مانند [۲۳] که در آن، یکی از این پیشنهادها بررسی شده است.

۴-۱. نمایه فرصت دیجیتال^۳ DOI

ITU نمایه فرصت فناوری اطلاعات و ارتباطات را به عنوان معیاری اصلی در سنجش جامعه اطلاعاتی از سال ۲۰۰۶ ایجاد کرده است [۲۴]. این معیار ترکیبی، امکان بهره بردن شهروندان یک کشور را از دسترسی به اطلاعات از نظر فراگیر بودن، موجود بودن در همه جا، برابر تقسیم شدن و کم هزینه بودن می‌سنجد. این نمایه در سه زمینه، کشورها را ارزیابی می‌کند:

1. Indicators
2. Index
3. Digital Opportunity index

۱. فرصت: دارای نشانگرهای زیر است:

- الف) درصد جمعیت زیر پوشش تلفن همراه؛
 ب) تعرفه دسترسی به اینترنت به صورت درصدی از درآمد هر فرد؛
 ج) تعرفه تلفن همراه به صورت درصدی از درآمد هر فرد.

۲. زیرساخت: دارای نشانگرهای زیر است:

- الف) نسبت خانواده‌های با یک خط تلفن ثابت به همه خانواده‌ها؛
 ب) نسبت خانواده‌های با یک کامپیوتر به همه خانواده‌ها؛
 ج) نسبت خانواده‌های دارای دسترسی به اینترنت در خانه به همه خانواده‌ها؛
 د) تعداد مشترکان تلفن همراه به ازای هر ۱۰۰ نفر مقیم؛

ه) تعداد مشترکان اینترنت به ازای هر ۱۰۰ نفر مقیم.

۳. بهره‌وری: دارای نشانگرهای زیر است:

- الف) نسبت افرادی که از اینترنت استفاده می‌کنند به همه افراد جامعه؛
 ب) نسبت مشترکان اینترنت با پهنای باند زیاد به کل مشترکان اینترنت؛
 ج) نسبت مشترکان تلفن همراه با پهنای باند زیاد به کل مشترکان تلفن همراه.

۴-۲. نمایه فناوری اطلاعات از UNCTAD^۴

این نمایه در سال ۱۹۶۴ پایه‌ریزی شد و در واقع، بخش مهمی از مجمع سازمان ملل متحد^۵ است که یکی از پنج سازمان اصلی زیرمجموعه سازمان ملل متحد است و با موضوع‌هایی همچون تجارت، سرمایه‌گذاری و توسعه سروکار دارد [۲۵]. بخش تحلیل فناوری اطلاعات، یکی از بخش‌های UNCTAD است که از سال ۲۰۰۴ به گردآوری اطلاعات و داده‌ها در زمینه فناوری اطلاعات پرداخته است؛ البته داده‌هایی از سال ۱۹۹۵ را نیز درون گزارش‌های خود دارد. گزارش سال ۲۰۱۰ نیز تعداد ۴۶

4. United Nations Conference on Trade and Development
 5. United Nations General Assembly

رتبه ۶۹ میان ۷۰ کشور، در ۲۰۰۸، رتبه ۷۰ میان ۷۰ کشور، در ۲۰۰۹، رتبه ۶۸ میان ۷۰ کشور و در ۲۰۱۰، رتبه ۶۹ میان ۷۰ کشور را داشته است. این آمارها بر پایه گزارش‌های سالیانه شرکت IBM و گروه اقتصادی هوشمند^۱ EIU است.

۶. میزان دسترسی به خدمات پشتیبانی کسب و کار الکترونیکی (ضریب ۰/۰۵).

۴-۴. نمایه فرصت‌های دیجیتال DAI

نمایه دسترسی دیجیتال به عنوان نخستین نمایه جهانی در رتبه‌بندی کشورها از منظر ICT در سال ۲۰۰۲، ۱۷۸ کشور جهان را در آمار خود بررسی کرده است؛ این رتبه‌بندی توسط اتحادیه بین‌المللی مخابرات ITU شده است. نمایه DAI ابزار سودمندی برای شناسایی آینده پیشرفت کشورهای جهان در امر ICT تلقی می‌شود، و وجه تمایز آن با دیگر نمایه‌های همانند، دارا بودن متغیرهای جدیدی همچون سطح تحصیلات و توانایی مالی است. DAI محدودیت‌های نمایه‌های پیشین را با تمرکز بر دسترسی، پوشش کشوری و انتخاب نشانگرهای مناسب برطرف می‌کند و ۸ نشانگر دارد که در ۵ دسته برای امتیازدهی کشورها گذاشته شده است. ملاک ارزیابی هر کشور، امتیازی است که در عددی میان صفر و ۱ به آن داده می‌شود. در نهایت، این امتیازها کشورها را در ۴ گروه عالی، بالا، متوسط و پایین دسته‌بندی می‌کند. نشانگرها در ادامه نوشته شده است.

۱. زیرساخت: تعداد مشترکان تلفن ثابت و همراه به ازای هر ۱۰۰ نفر جمعیت؛
۲. توانایی مالی: نرخ دسترسی به اینترنت به عنوان درصدی از درآمد سرانه؛
۳. دانش: میزان سواد در افراد بالغ و میزان ثبت‌نام برای سطح‌های اول، دوم و سوم مدرسه‌ها؛
۴. کیفیت: پهنای باند اینترنت بین‌المللی به جمعیت و مشترکان اینترنت پهن‌بند به ازای هر ۱۰۰ نفر؛
۵. به کارگیری: تعداد کاربران اینترنت به ازای هر ۱۰۰ نفر جمعیت.

نشانگر را تعریف می‌کند که ۵ نشانگر بیشتر نسبت به گزارش پیشین در سال ۲۰۰۵ دارد [۲۶]. نشانگرها در ۶ دسته زیر گذاشته شده‌اند:

۱. نشانگرهای زیرساخت؛
۲. نشانگرهای دسترسی و به کارگیری ICT به وسیله افراد و خانواده‌ها؛
۳. نشانگرهای به کارگیری ICT به وسیله بخش‌های اقتصادی؛
۴. نشانگرهای بخش‌های ساخت و گسترش ICT؛
۵. نشانگرهای تجارت جهانی کالاهای ICT؛
۶. نشانگرهای ICT در آموزش.

۴-۳. آمادگی الکترونیکی (e-readiness)

اتحادیه‌ای از چند شرکت بزرگ در زمینه فناوری اطلاعات، این نمایه را ارائه کرده است. شرکت IBM یکی از شرکت‌های درون این اتحادیه است. این اتحادیه از سال ۲۰۰۰ تاکنون، میزان آمادگی الکترونیکی ۶۸ اقتصاد برتر دنیا را بر پایه همین نمایه بررسی کرده و گزارش آن را سالیانه منتشر نموده است. این بررسی در قالب ۱۰۰ نشانگر کمی و کیفی و در ۶ دسته کلی زیر انجام می‌شود [۲۷]:

۱. اتصال و زیرساخت‌های فناوری (ضریب ۰/۲۵)؛
۲. شرایط اقتصادی (ضریب ۰/۲۰)؛
۳. میزان ارتباط میان مصرف‌کننده و تولیدکننده فرآورده‌های ICT (ضریب ۰/۲۰)؛
۴. شرایط حقوقی و سیاسی (ضریب ۰/۱۵)؛
۵. شرایط اجتماعی و فرهنگی (ضریب ۰/۱۵)؛

البته تعریف‌های هر کدام از دسته‌ها و ضریب‌هایشان، تاکنون تغییرهایی داشته است. امتیازی که به هر کشور داده می‌شود، امتیازی بین صفر تا ۱۰ است که رتبه‌بندی نیز بر پایه آن شکل می‌گیرد. برای محاسبه امتیاز نهایی، روش میانگین وزنی به کار برده شده است. وزن هر کدام از معیارها در روبه‌روی آن‌ها نوشته شده است. ایران در ۲۰۰۱، رتبه ۵۰ میان ۶۰ کشور، در ۲۰۰۲، رتبه ۵۳ میان ۶۰ کشور، در ۲۰۰۳، رتبه ۵۲ میان ۶۰ کشور، در ۲۰۰۴، رتبه ۵۷ میان ۶۴ کشور، در ۲۰۰۵، رتبه ۵۹ میان ۶۵ کشور، در ۲۰۰۶، رتبه ۶۵ میان ۶۸ کشور، در ۲۰۰۷،

۴-۵. نمایه فرصت فناوری اطلاعات و ارتباطات ICT-OI

این نمایه، یک چارچوب مفهومی سه سطحی فراهم می‌کند

که در زیرسطح‌های این چارچوب مفهومی نوشته شده است:

۱. آمادگی ICT: نشان‌دهنده سطح زیرساخت‌های شبکه‌ای و دسترسی به ICT؛

۲. شدت ICT: نشان‌دهنده به کارگیری ICT در جامعه؛

۳. تأثیر ICT: نشان‌دهنده نتیجه‌ها و خروجی‌های مؤثر در

به کارگیری ICT.

بر پایه این چارچوب مفهومی، سه زیرنمایه «دسترسی»، «به کارگیری» و «مهارت» برای IDI تعریف شده است. در زیر، نشانگرهای هر کدام از این سه زیرنمایه به همراه درصد تأثیر آنها در محاسبه نهایی IDI نوشته شده است:

۱. دسترسی به ICT (چهل درصد)

الف) تعداد تلفن‌های ثابت به ازای هر ۱۰۰ نفر سکنه (%/۲۰)

ب) تعداد اشتراک‌های تلفن همراه به ازای هر ۱۰۰ نفر سکنه (%/۲۰)؛

ج) پهنای باند اینترنت بین‌المللی بر حسب بیت بر ثانیه به ازای هر کاربر اینترنت (%/۲۰)؛

د) نسبت خانواده‌های دارنده رایانه به همه خانواده‌ها (%/۲۰)؛

ه) نسبت خانواده‌های دارنده دسترسی به اینترنت به همه خانواده‌ها (%/۲۰).

۲. به کارگیری ICT (چهل درصد)

الف) تعداد کاربران اینترنت به ازای هر ۱۰۰ سکنه (%/۳۳)؛

ب) تعداد مشترکان باندپهن به ازای هر ۱۰۰ سکنه (%/۳۳)؛

ج) تعداد مشترکان تلفن همراه به ازای هر ۱۰۰ سکنه (%/۳۳).

۳. مهارت‌های ICT (بیست درصد)

الف) نرخ سواد بزرگسالان (%/۳۳)؛

ب) نسبت ثبت‌نام در دومین مقطع تحصیلی در یک سال تحصیلی (%/۳۳)؛

ITU نمایه فرصت فناوری اطلاعات و ارتباطات را به عنوان معیاری اصلی در سنجش جامعه اطلاعاتی از سال ۲۰۰۵ ایجاد کرده است. این نمایه، مدلی مفهومی است که نحوه دسترسی به اینترنت و کاربرد آن را می‌سنجد و در دو زمینه زیر، کشورها را ارزیابی می‌کند:

۱. چگالی اطلاعات: به توانایی بهره‌ورانه و ظرفیت اقتصادی و مفهوم‌هایی همچون کارمندان ICT و سرمایه ICT می‌پردازد و دارای نشانگرهای زیر است:

الف) شبکه: تعداد خط‌های تلفن در هر صد نفر، مشترکان تلفن همراه در هر صد نفر، پهنای باند بین‌المللی؛

ب) مهارت: نسبت سواد بزرگسالان، میزان ثبت‌نام مدرسه‌ها.

۲. کاربرد اطلاعات: به کاربرد اقتصاد یک کشور از ICT

اشاره دارد و دارای نشانگرهای زیر است:

الف) کاربرد: کاربران اینترنت در هر صد نفر، نسبت خانواده‌های دارای تلویزیون، و تعداد رایانه در هر صد نفر؛

ب) چگالی: تعداد مشترکان باندپهن اینترنت در هر صد نفر، تعرفه مکالمه‌های بین‌المللی خروجی (دقیقه).

۴-۶. نمایه توسعه فناوری اطلاعات و ارتباطات IDI^۱

این نمایه جایگزین نمایه‌های DAI, DOI و ICT-OI شده است. این جایگزینی به درخواست عضوهای ITU انجام شده است که با ترکیب کردن نمایه‌های پیشین، نمایه یکتایی برای دنبال کردن شکاف دیجیتال و معیاری برای توسعه جامعه اطلاعاتی ارائه می‌دهد [۲۸]. هدف‌های اصلی IDI اندازه‌گیری موردهای زیر است [۲۹]:

۱. سطح توسعه ICT و تکوین آن در گذر زمان در کشورها؛

۲. پیشرفت ICT در کشورهای توسعه‌یافته و در حال توسعه؛

۳. شکاف دیجیتال؛

۴. توانایی بالقوه پیشرفت در ICT بر پایه توانایی‌ها و

مهارت‌ها؛

1. ICT Development Index

۶. داده‌های گردآوری شده

هر کدام از جدول‌های این بخش، از منابع‌های گوناگونی گردآوری شده که به دست آوردن برخی از آنها بسیار دشوار یا دست‌کم زمان‌گیر بوده است. نگارندگان کوشش کرده‌اند که اطلاعات مربوط به هم را در یک جدول قرار دهند. برخی از این اطلاعات، از سایت آمار ایران و برخی نیز از سایت‌های شرکت مخابرات ایران و دیگر منابع گردآوری شده است. این داده‌ها تا سال ۱۳۸۸ گردآوری شده و بدین سبب، برخی از داده‌ها با وضعیت کنونی متفاوت است؛ البته استان‌هایی مانند خراسان رضوی و همانند آن نیز، که در آن زمان به تازگی از هم جدا شده بودند، به دلیل اینکه تازه افزوده شده بودند، داده‌های کافی نداشتند.

۱-۶. به کارگیری

به کارگیری فناوری اطلاعات در یک استان، از معیارهای مهم به دست آوردن شکاف دیجیتال است. جدول (۱) دربردارنده تعداد شعبه‌های دارنده بانکداری الکترونیک^۱ است که برای سادگی، «بانک» نامیده شده است؛ تعداد اداره‌های ICT روستایی^۲، ستون دیگری از این جدول است که «روستا» نامیده شده است؛ ستون‌های دیگر جدول به ترتیب عبارت‌اند از: تعداد مدرسه‌های متصل به شبکه^۳ که «مدرسه» نامیده شده؛ تعداد دانشگاه‌های متصل به شبکه جهانی که «دانشگاه» نامیده شده و تعداد کاربران اینترنت^۴ که «اینترنت» نامیده شده است.

(ج) نسبت ثبت‌نام در سومین مقطع تحصیلی در یک سال تحصیلی (۳۳٪).

بر پایه این نمایه، رتبه ایران در سال ۲۰۰۷ برابر با ۸۶ و در سال ۲۰۰۸ برابر با ۸۴ بوده است.

۵. شاخص‌های شکاف دیجیتال به کار گرفته شده در این پژوهش

در این پژوهش، سعی شده با توجه به مقالات و نشانگرهای ارائه شده و هماهنگ کردن آن‌ها با وضعیت کنونی کشور، بهترین نشانگرها برگزیده شود و با گردآوری آمار مربوط به آن‌ها، اندازه گسترش فناوری اطلاعات استان‌ها ارزیابی و وضعیت فاصله دیجیتالی میان آن‌ها بررسی گردد؛ بنابراین، نشانگرهای برگزیده شده در سه گروه زیر دسته‌بندی شد:

الف) زیرساخت‌ها: نرخ نفوذ تلفن ثابت، نرخ نفوذ تلفن همراه، وضعیت تلفن همگانی، وضعیت شبکه داده (تعداد پورت‌های دایر شده)؛

ب) به کارگیری: تعداد مرکزهای ICT روستایی (بهره‌برداری شده)، بانکداری الکترونیک (تعداد شعبه)، وضعیت مدرسه‌ها (اتصال به اینترنت)، وضعیت دانشگاه‌ها (اتصال به اینترنت با فیبر نوری)، تعداد کاربران اینترنت؛

ج) منابع انسانی: درصد باسوادان (شهری-روستایی)، تعداد دانشجویان، سطح درآمد، جمعیت (شهری-روستایی)، تعداد خانواده‌های دارای رایانه شخصی.

جدول (۱): داده‌های به کارگیری

ردیف	استان	بانک	روستا	مدرسه	دانشگاه	اینترنت
۱	آذربایجان شرقی	۷۰۰	۷۵۰	۳۷۵	۱۰	۲۷.۵
۲	آذربایجان غربی	۴۰۷	۲۵۶	۵۳	۱۹	۱۹.۳
۳	اردبیل	۳۱۲	۲۵۶	۷۳	۴	۱۸.۳
۴	اصفهان	۹۵۶	۳۲۰	۳۳۵	۱۲	۲۷.۵
۵	ایلام	۲۳۹	۶۷	۱۵۱	۴	۲۴.۵

1. E banking, number of branch
2. number of office in ICT village
3. number of school connected to network
4. number of users of internet or internet penetration

ردیف	استان	بانک	روستا	مدرسه	دانشگاه	اینترنت
۶	بوشهر	۳۶۹	۱۷۲	۳۱۶	۴	۲۷.۵
۷	تهران	۳۸۷۲	۲۱۶	۱۳۴۹	۷۶	۲۷.۵
۸	چهارمحال و بختیاری	۳۰۳	۱۱۱	۲۰۳	۴	۱۷.۳
۹	خراسان جنوبی	۱۸۲	۱۱۳	۸۸	۱۰	۲۷.۵
۱۰	خراسان رضوی	۱۲۸۲	۴۶۶	۵۹۸	۱۵	۲۷.۵
۱۱	خراسان شمالی	۱۶۵	۷۵	۱۰۳	۴	۲۷.۵
۱۲	خوزستان	۷۳۸	۱۴۶	۵۰	۱۸	۲۱.۶
۱۳	زنجان	۲۸۵	۱۵۸	۵۸	۱۰	۲۳.۷
۱۴	سمنان	۳۲۸	۸۷	۳۳۷	۱۲	۲۷.۵
۱۵	سیستان و بلوچستان	۳۱۰	۱۵۷	۱۰۶	۲۵	۲۷.۱
۱۶	فارس	۱۱۲۵	۳۱۷	۳۷۱	۱۲	۲۷.۵
۱۷	قزوین	۳۳۱	۱۵۶	۱۲۶	۴	۲۷.۵
۱۸	قم	۳۲۱	۵۰	۶۸	۷	۲۷.۵
۱۹	کردستان	۲۴۰	۱۲۵	۶۲	۴	۱۲.۳
۲۰	کرمان	۵۱۶	۲۰۰	۱۰۶	۸	۱۳
۲۱	کرمانشاه	۴۴۸	۱۴۰	۱۳۹	۵	۱۶.۲
۲۲	کهگیلویه و بویراحمد	۱۱۸	۱۵۳	۹۵	۴	۸.۶
۲۳	گلستان	۳۳۶	۲۲۰	۹۷	۴	۲۷.۵
۲۴	گیلان	۴۱۳	۲۵۲	۸۸	۷	۱۵
۲۵	لرستان	۲۷۵	۲۱۲	۶۰	۴	۹.۴
۲۶	مازندران	۶۶۸	۲۸۴	۲۱	۱۶	۲۲.۴
۲۷	مرکزی	۳۱۰	۲۵۷	۶۴	۷	۲۱.۶
۲۸	هرمزگان	۳۴۲	۱۹۹	۸۲	۸	۲۴.۹
۲۹	همدان	۴۰۵	۲۶۸	۷۰	۵	۲۰.۸
۳۰	یزد	۳۸۴	۱۲۷	۲۰۰	۵	۲۷.۵

۲-۶. زیرساخت

در بسیاری از روش‌های آن به کار گرفته می‌شود، ترافیک شبکه برای تلفن و مهم‌تر از آن برای داده است، ولی این داده با همه پیگیری‌های انجام‌شده، در اختیار این پژوهش قرار نگرفت. داده‌های گردآوری شده شامل ضریب نفوذ تلفن همراه^۲، ضریب نفوذ تلفن ثابت^۳ و تعداد تلفن عمومی^۴ در هر استان است.

جدول (۲) نشان‌دهنده اطلاعات مربوط به زیرساخت‌های فناوری اطلاعات است. این داده‌ها به کمک آمارهای چندساله بر روی سایت‌های شرکت مخابرات، گردآوری شده است. مهم‌ترین داده‌ای که برای معیارهای شکاف دیجیتال اهمیت فراوانی دارد و

2. mobile penetration
3. fixed line penetration
4. public phone

1. Infrastructure

جدول (۲): زیرساخت

ردیف	استان	ضریب نفوذ تلفن همراه	ضریب نفوذ تلفن ثابت	تعداد تلفن‌های عمومی
۱	آذربایجان شرقی	۲۸.۴۰	۳۸.۰۳	۱۱۲۰۸
۲	آذربایجان غربی	۲۷.۲۹	۳۲.۶۷	۸۱۱۰
۳	اردبیل	۲۴.۷۱	۳۳.۶۳	۴۶۳۱
۴	اصفهان	۴۲.۰۳	۳۹.۱۷	۱۴۱۵۰
۵	ایلام	۲۹.۳۲	۲۵.۲۷	۱۵۳۵
۶	بوشهر	۳۹.۰۳	۳۴.۵۳	۲۹۷۹
۷	تهران	۴۸.۳۸	۴۳.۴۲	۳۵۰۸۰
۸	چهارمحال و بختیاری	۲۹.۹۱	۲۷.۶۸	۱۷۱۵
۹	خراسان جنوبی	۲۱.۰۷	۲۹.۹۴	۱۷۷۹
۱۰	خراسان رضوی	۲۵.۷۱	۳۰.۴۴	۱۳۷۳۲
۱۱	خراسان شمالی	۱۷.۱۹	۲۵.۴۵	۱۵۵۴
۱۲	خوزستان	۳۳.۱۸	۲۱.۸۶	۱۲۱۹۹
۱۳	زنجان	۲۹.۶۷	۲۹.۶۷	۲۸۲۱
۱۴	سمنان	۴۲.۲۵	۴۱.۷۱	۲۶۱۶
۱۵	سیستان و بلوچستان	۱۸.۲۵	۱۵.۸۷	۵۴۳۰
۱۶	فارس	۳۹.۸۹	۳۱.۸۵	۱۳۴۲۴
۱۷	قزوین	۳۴.۰۴	۳۲.۸	۳۵۱۸
۱۸	قم	۴۱.۵۹	۳۴.۰۴	۳۶۳۷
۱۹	کردستان	۲۵.۷۲	۲۸.۵۳	۳۳۶۵
۲۰	کرمان	۲۶.۷۲	۲۵.۶۸	۶۲۸۳
۲۱	کرمانشاه	۳۱.۵۳	۲۵.۹۶	۵۴۵۰
۲۲	کهگیلویه و بویراحمد	۲۳.۲۳	۱۹.۸۷	۱۵۲۵
۲۳	گلستان	۲۴.۹۷	۳۰.۱۴	۳۷۴۵
۲۴	گیلان	۳۳.۱۰	۳۶.۲۹	۷۳۱۸
۲۵	لرستان	۲۲.۹۲	۲۳.۹	۳۷۸۷
۲۶	مازندران	۳۵.۱۱	۴۴.۹	۱۰۲۷۱
۲۷	مرکزی	۲۸.۰۹	۳۲.۴۶	۴۲۸۹
۲۸	هرمزگان	۳۸.۱۳	۲۴.۹۵	۲۴۷۶
۲۹	همدان	۲۵.۹۴	۳۱.۴۷	۳۹۲۲
۳۰	یزد	۳۶.۷۴	۳۸.۵۶	۳۲۸۸

۳-۶. منابع انسانی

یکی از مهم‌ترین بخش‌های بررسی شکاف دیجیتال، نشانگرهای مربوط به انسان‌هاست. در این بخش، تعدادی از داده‌های یافت‌شده، درون جدول (۳) گذاشته شده است. جمعیت شهری^۱ یک ستون از این جدول است که برای سادگی، «شهری»

نامیده شده است. ستون‌های دیگر جدول به ترتیب عبارت‌اند از: جمعیت روستایی^۲ که «روستایی» نامیده شده؛ درآمد^۳ که بر حسب ۱۰۰ هزار ریال درآمد خانواده نوشته شده؛ تعداد دانشجو که «دانشجو» نامیده شده؛ تعداد رایانه شخصی که «رایانه» نامیده شده؛

2. Rural population
3. Income

1. city population

درصد باسوادان شهری که «با شهری» نامیده شده، و درصد باسوادان روستایی که «با روستا» نامیده شده است.

جدول (۳): منابع انسانی

ردیف	استان	شهری	روستایی	درآمد	دانشجو	رایانه	با شهر	با روستا
۱	آذربایجان شرقی	۲۴۰۲۵۳۹	۱۲۰۰۸۲۰	۵۲۱.۳	۱۱۷۸۰۴	۱۵۸۹۱۷	۶۰	۲۴
۲	آذربایجان غربی	۱۷۲۴۹۵۴	۱۱۴۸۵۰۵	۶۰۵.۰۳	۴۹۱۶۰	۸۷۰۵۶	۵۲	۲۷
۳	اردبیل	۷۱۵۵۹۷	۵۱۲۱۹۵	۵۸۰.۹۸	۳۰۱۵۴	۳۸۵۹۴	۵۰	۳۰
۴	اصفهان	۳۷۹۸۱۷۲۸	۷۵۸۸۹۰	۵۷۶.۵۳	۱۴۳۳۷۶	۳۳۰۰۱۲	۷۴	۱۳
۵	ایلام	۳۳۱۲۳۱	۲۱۰۷۰۳	۶۰۶	۱۵۸۲۸	۱۵۰۶۷	۵۲	۲۹
۶	بوشهر	۵۷۷۴۶۵	۳۰۳۴۰۹	۶۰۰.۶۸	۱۹۰۴۸	۴۳۴۱۱	۵۸	۲۸
۷	تهران	۱۲۲۶۰۴۳۱	۱۱۶۱۸۸۹	۷۱۶.۵	۴۹۷۵۹۹	۱۳۸۱۴۷۹	۸۴	۶
۸	چهارمحال و بختیاری	۴۴۲۲۹۸	۴۱۴۶۲۴	۵۶۱.۶۵	۲۱۵۳۰	۳۳۸۵۵	۴۵	۳۷
۹	خراسان جنوبی	۳۲۶۶۹۵	۳۰۸۳۰۵	۳۸۴.۲	۱۷۷۱۶	۱۹۸۰۶	۴۶	۳۵
۱۰	خراسان رضوی	۳۸۱۱۹۰۰	۱۷۷۹۹۸۰	۱۱۹.۹۷	۱۲۰۴۶۷	۲۶۱۰۴۱	۶۲	۲۴
۱۱	خراسان شمالی	۳۹۲۴۵۸	۴۱۴۳۶۵	۴۸۲.۵	۱۴۶۹۹	۲۳۸۷۸	۴۲	۳۶
۱۲	خوزستان	۲۸۷۳۵۶۴	۱۳۸۳۹۴۶	۶۱۰.۱۷	۱۲۵۵۰۰	۱۸۴۲۶۲	۶۰	۲۴
۱۳	زنجان	۵۵۹۳۴۰	۴۰۵۲۶۱	۵۶۵.۷۳	۳۶۴۱۷	۳۵۷۸۸	۵۰	۳۰
۱۴	سمنان	۴۴۰۵۵۹	۱۴۹۱۸۳	۵۲۸.۳	۵۷۱۴۹	۳۸۵۵۰	۶۸	۲۰
۱۵	سیستان و بلوچستان	۱۱۹۳۱۹۸	۱۲۰۶۵۴۷	۴۶۹.۴۱	۶۰۸۴۰	۴۵۲۷۶	۳۸	۳۰
۱۶	فارس	۲۶۵۲۹۴۷	۱۶۵۰۶۱۴	۶۳۲.۸۱	۱۲۱۲۵۳	۲۳۵۹۵۶	۵۵	۳۰
۱۷	قزوین	۷۷۷۹۷۵	۳۶۵۲۰۳	۶۴۱.۸	۵۶۹۸۴	۵۵۷۰۵	۶۰	۲۵
۱۸	قم	۹۸۳۰۹۴	۶۳۶۳۹	۴۹۷.۸	۲۰۹۳۷	۶۵۳۷۷	۸۲	۴
۱۹	کردستان	۸۵۵۸۱۹	۵۸۴۳۳۷	۵۴۲.۹۶	۱۸۴۸۲	۴۳۹۹۳	۵۰	۲۸
۲۰	کرمان	۱۵۵۲۵۱۹	۱۰۸۹۷۴۸	۵۳۳.۶۸	۷۲۲۷۹	۹۴۳۹۴	۵۲	۳۰
۲۱	کرمانشاه	۱۲۵۵۳۱۹	۶۱۸۷۱۸	۵۴۰.۴۲	۳۳۱۰۹	۶۱۲۵۸	۵۷	۲۵
۲۲	کهگیلویه و بویراحمد	۳۰۲۱۹۲	۳۲۹۸۴۹	۶۸۰	۲۲۵۰۵	۱۷۲۴۵	۴۳	۳۸
۲۳	گلستان	۷۹۵۱۲۶	۸۱۹۵۸۴	۵۲۶	۳۲۴۵۷	۵۴۲۹۱	۴۳	۳۸
۲۴	گیلان	۱۲۹۵۷۵۱	۱۱۰۹۱۰۴	۵۹۲	۶۱۶۲۲	۱۰۲۱۱۷	۴۹	۳۴
۲۵	لرستان	۱۰۲۰۱۵۰	۶۹۱۴۴۸	۵۰۸	۵۴۱۰۹	۴۸۷۸۳	۵۱	۳۰
۲۶	مازندران	۱۵۵۴۱۴۳	۱۳۶۸۲۳۳	۶۸۵.۵	۱۲۳۵۲۵	۱۳۹۰۱۲	۴۸	۳۷
۲۷	مرکزی	۹۳۲۰۷۳	۴۱۹۱۸۴	۵۲۱	۷۴۸۸۷	۷۲۸۹۹	۶۱	۲۳
۲۸	هرمزگان	۶۶۱۳۲۵	۷۴۰۶۰۵	۶۴۶	۲۵۴۲۷	۴۹۹۸۰	۴۲	۴۰
۲۹	همدان	۹۸۰۷۷۱	۷۲۱۲۲۵	۵۲۲.۳	۴۴۰۱۰	۶۳۹۲۳	۵۰	۳۲
۳۰	یزد	۷۸۹۸۰۳	۲۰۰۹۸۸	۵۴۲.۳۵	۵۵۳۶۷	۶۴۲۵۶	۷۲	۱۶

۷. خوشه‌بندی

تعداد داده‌های بیشتر نباشد. الگوریتم K-means یک الگوریتم تکرارشونده است که با یک بخش‌بندی تصادفی اولیه آغاز می‌شود و جابجا کردن داده‌ها میان خوشه‌ها را ادامه می‌دهد تا مجموع مربعات خطاهای درون خوشه‌ها کمینه شود [۳۳]. این الگوریتم دربردارنده چهار مرحله است [۳۴]:

۱. تعداد K مرکز خوشه از میان داده‌ها به صورت تصادفی برگزیده می‌شود (روش‌های گوناگونی برای بهبود این گزینش اولیه ارائه شده است).

۲. دیگر داده‌های باقی‌مانده درون خوشه‌ای گذاشته می‌شوند که کمترین فاصله را از مرکز آن دارند.

۳. بر پایه داده‌های کنونی درون هر خوشه، یک مرکز خوشه جدید محاسبه می‌شود.

۴. تابع کیفیت یا ارزیابی محاسبه می‌شود و اگر نتیجه هنوز خوب نبود، به مرحله ۲ برمی‌گردد و گرنه خوشه‌های مناسب یافت می‌شود و داده‌ها میان آن‌ها تقسیم می‌شوند.

الگوریتم K-means می‌کوشد تا تابع زیر را که مجموع مربعات خطاهاست، کمینه کند.

$$e^2 = \sum_{j=1}^k \sum_{i=1}^{n_j} \|x_i^{(j)} - y_j\|^2$$

در این تابع، $x_i^{(j)}$ برابر i امین داده‌ای است که درون خوشه j ام قرار دارد و y_j مرکز خوشه j ام است. تاکنون شکل‌های گوناگونی از این الگوریتم پیشنهاد و پیاده‌سازی شده است تا کار خوشه‌بندی به کمک آن را بهبود بخشد.

۷-۱. ابزار به کار گرفته شده

ابزارهای آماده گوناگونی برای خوشه‌بندی وجود دارد. در این پژوهش، از کلمن‌تاین^۴ استفاده شده که برای به کارگیری آن باید SPSS نصب باشد؛ البته می‌توان به جای SPSS ابزار متن باز PSPP را نیز به کار برد. SPSS یک نرم‌افزار آماری است و نسخه متن باز آن PSPP نیز برخی از امکانات SPSS را

خوشه‌بندی^۱ یکی از شاخه‌های داده‌کاوی^۲ است و به فرآیند تقسیم مجموعه‌ای از داده‌ها (یا شیء‌ها) به زیردسته‌هایی با مفهوم خوشه گفته می‌شود. به این ترتیب یک خوشه، یک سری داده‌های همانند است که مانند یک گروه واحد رفتار می‌کند. خوشه‌بندی همان دسته‌بندی^۳ است، با این تفاوت که دسته‌ها از پیش تعریف شده و معین نمی‌باشند و گروه‌بندی داده‌ها بدون نظارت انجام می‌گیرد.

فرض کنید که مجموعه داده‌های X مورد نظر ما از نقطه‌های داده‌ای (یا الگوها، گروه‌ها یا رکوردها)، در فضای ویژگی A تشکیل شده باشد، در این صورت، هدف خوشه‌بندی، یافتن زیرمجموعه‌هایی است که اجتماع همه آن‌ها کل مجموعه بوده و دو به دو آن‌ها اشتراکی نداشته باشند.

فرمول ۱:

$$X = C_1 \cup C_2 \cup \dots \cup C_k \quad \text{و} \quad C_1 \cap C_2 \cap \dots \cap C_k = \phi$$

شیء‌های خوشه‌بندی شده بر پایه اصل بیشترین شباهت میان عضوهای هر دسته و کمترین شباهت میان دسته‌های گوناگون گروه‌بندی می‌شود، یعنی خوشه‌ها به گونه‌ای تنظیم می‌شود که شیء‌های درون هر خوشه، بیشترین شباهت را با یکدیگر داشته باشد. هر خوشه به عنوان یک کلاس است که قانون‌هایی از آن مشتق می‌شود. می‌توان خوشه‌بندی را به صورت دسته‌بندی تعریف کرد، با این تفاوت که کلاس‌ها و برچسب آن‌ها از پیش تعریف شده نیست و خوشه‌بندی بدون نظارت انجام می‌گیرد.

روش خوشه‌بندی k-means، یکی از روش‌های مهم و به نسبت قدیمی در خوشه‌بندی داده‌هاست که امروزه نیز به خوبی کاربرد دارد [۳۰-۳۲] و داده‌ها را به k گروه مجزا بخش می‌کند. در این الگوریتم باید در آغاز، تعداد K گروه مشخص باشد، سپس خوشه‌بندی انجام شود؛ البته این تعداد باید از

1. Clustering
2. Data mining
3. Classification

4. Clementine

دارد. کلمن تاین کمک می کند تا به سادگی بتوان الگوریتم های گوناگون خوشه بندی و داده کاوی را بر روی داده ها به کار برد.

۲-۷. خوشه بندی استان های کشور بر پایه شکاف دیجیتالی

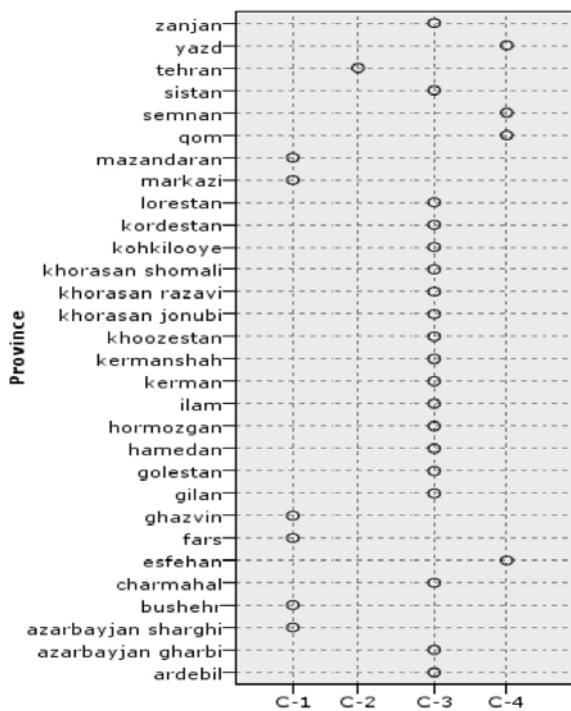
در این بخش، به نتیجه های خوشه بندی داده های استان های گوناگون کشور که پیش از این نوشته شد، پرداخته می شود. روش k-means برای خوشه بندی به کار برده شده است. تعداد خوشه ها در این الگوریتم باید از پیش مشخص شده باشد تا الگوریتم بتواند بر پایه آن خوشه بندی را انجام دهد. راه های گوناگونی برای به دست آوردن تعداد خوشه ها وجود دارد. یکی از این راه ها به کارگیری الگوریتم خوشه بندی two-step است، اما به کارگیری آن بر روی این داده ها فقط باعث شناسایی دو خوشه شد که در یک خوشه، تهران و در خوشه دوم، استان های دیگر قرار گرفت. از سوی دیگر در این کار، به دست آوردن اینکه چه استان هایی با توجه به تعداد خوشه گوناگون در کنار یکدیگر گذاشته می شوند، سودمند است؛ بنابراین، چندین بار الگوریتم k-means با تعداد خوشه متفاوت (از تعداد خوشه ۲ تا تعداد خوشه ۴) اجرا شد. برای تعداد خوشه دو، نتیجه همانند الگوریتم two-step است و تهران در یک خوشه و دیگر استان ها در خوشه دیگر گذاشته می شوند، اما برای تعداد خوشه سه، نتیجه خوشه بندی به قرار زیر است:

خوشه یکم: آذربایجان شرقی، اصفهان، بوشهر، فارس، قزوین، مرکزی، قم، سمنان و یزد؛

خوشه دوم: تهران؛

خوشه سوم: دیگر استان ها.

شکل (۱)، خروجی کلمن تاین را برای الگوریتم k-means برای خوشه بندی با تعداد خوشه چهار نشان می دهد؛ البته برای اینکه این شکل بهتر در صفحه نشان داده شود، فاصله میان خوشه ها (فاصله میان خط های عمودی) در شکل کم شده است.



شکل (۱): نمودار خروجی الگوریتم K-MEANS برای همه نشانگرها

در این خوشه بندی، هر کدام از خوشه ها دارای استان های زیرند:

خوشه یکم: آذربایجان شرقی، بوشهر، فارس، قزوین، مرکزی، مازندران؛

خوشه دوم: تهران؛

خوشه سوم: اردبیل، آذربایجان غربی، چهار محال و بختیاری، گیلان، گلستان، همدان، هرمزگان، ایلام، کرمان، کرمانشاه، خوزستان، خراسان جنوبی، خراسان رضوی، خراسان شمالی، کهگیلویه و بویر احمد، کردستان، لرستان، سیستان و بلوچستان، زنجان؛

خوشه چهارم: اصفهان، قم، سمنان، یزد.

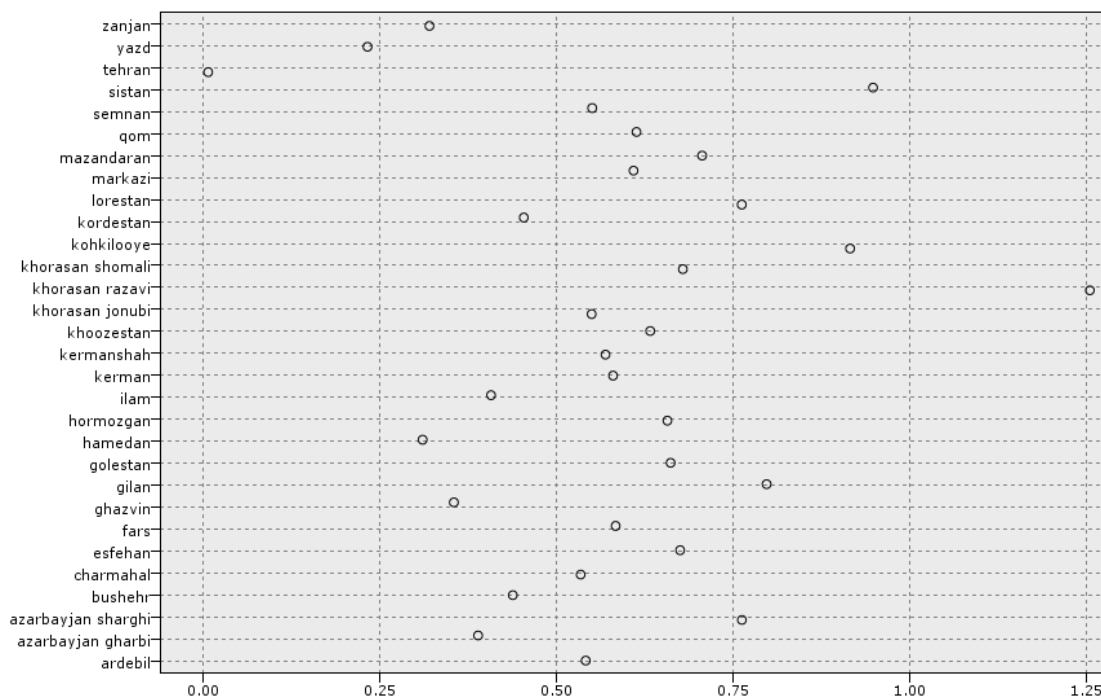
این شکل به خوبی، تفاوت میان تهران را با دیگر استان ها نشان می دهد، زیرا باز هم تهران در یک خوشه به صورت جداگانه گذاشته شده و در بهترین وضعیت از نظر نشانگرهای فناوری اطلاعات است و بدترین وضعیت را استان های خوشه سوم دارند. از این نمودار، روشن است که امکانات به صورت ناعادلانه ای میان استان های کشور پخش شده است؛ به گونه ای که تهران در همه نشانگرها آن چنان تفاوتی با دیگر استان ها دارد که همواره در یک خوشه جداگانه قرار می گیرد و در بیشتر زمینه ها بسیار بالاتر

شکل (۲)، فاصله میان هر استان را با مرکز خوشه خود در خوشه‌بندی با تعداد خوشه چهار نشان می‌دهد. به این ترتیب تقریباً روشن می‌شود که هر کدام از استان‌ها چه فاصله‌ای با عددهای داده شده در جدول (۴) دارند و بهتر می‌توان به مقایسه میان استان‌ها پرداخت.

از دیگر استان‌هاست و امکانات بسیار بیشتری در اختیار آن گذاشته شده است. برای مقایسه میان دیگر استان‌ها، جدول (۴) کمک‌کننده است. این جدول، میانگین برخی از نشانگرها را برای هر خوشه نشان می‌دهد و تا اندازه‌ای می‌توان به کمک آن، وضعیت شکاف دیجیتال میان دیگر استان‌ها را نیز شناسایی کرد.

جدول (۴): میانگین برخی از نشانگرهای هر خوشه

خوشه چهار	خوشه سه	خوشه دو	خوشه یک	نشانگر
۵۳۶/۲۴۵	۵۳۰/۳۶۸	۷۱۶/۵	۶۰۰/۵۱۵	درآمد
۴۹۷/۲۵	۳۸۵/۰۵۳	۳۸۷۲/۰	۵۸۳/۸۳۳	بانک
۳۸/۳۷	۲۷/۳۳	۴۳/۴۲	۳۵/۷۶۲	ضریب نفوذ تلفن ثابت
۲۷/۵	۲۰/۱۰۵	۲۷/۵	۲۵/۶۶۷	اینترنت
۴۰/۶۵۲	۲۶/۷۶۶	۴۸/۳۸	۳۴/۰۹۳	ضریب نفوذ تلفن همراه
۶۹۲۰۷/۲۵	۴۵۰/۶۹	۴۹۷۵۹۹	۸۵۵۸۳/۵	دانشجو
۱۲۴۵۴۸/۷۵	۶۷۴۰۰/۳۶۸	۱۳۸۱۴۷۹	۱۱۷۶۵۰	رایانه
۷۴	۴۹/۱۵۸	۸۴	۵۷	با شهر
۱۳/۲۵	۳۱/۴۲۱	۶	۲۷/۸۳۳	با روستا
۵۹۲۲/۷۵	۴۸۰۹/۳۱۶	۳۵۰/۸۰	۷۶۱۴/۸۳۳	تعداد تلفن‌های عمومی
۲۳۵	۱۲۰/۱۰۵	۱۳۴۹	۲۱۲/۱۶۷	مدرسه
۹	۸/۵۲۶	۷۶	۸/۸۳۳	دانشگاه
۱۴۶	۱۸۸/۱۰۵	۲۱۶	۳۲۲/۶۶۷	روستا



شکل (۲): فاصله میان هر استان با مرکز خوشه در خوشه‌بندی با تعداد خوشه چهار

۸. نتیجه گیری

مسئولان مقصرونند. در [۳۵] تعدادی از عامل‌های تشدیدکننده این شکاف توضیح داده شده است که به سادگی می‌توان با از میان برداشتن یا کاهش اثر آن‌ها، شکاف میان استان‌های کشور در زمینه فناوری اطلاعات را کاهش داد. برخلاف دیگر زمینه‌ها که تمرکززدایی می‌تواند هزینه بر باشد، تمرکززدایی در زمینه فناوری اطلاعات هزینه چندانی ندارد و بسیار ساده‌تر می‌تواند انجام شود، زیرا نیاز به ساخت کارخانه‌ها و صنایع سنگین در آن وجود ندارد و ابزار لازم برای آن، رایانه و خط اینترنت با سرعت مناسب است. امید است با برنامه‌ریزی بهتر، این شکاف کاهش یابد.

این پژوهش برای کامل‌تر شدن، نیاز به پژوهش‌های گوناگون دیگری نیز دارد تا بتوان به نتیجه‌های دقیق‌تری در این زمینه رسید. یکی از کارهایی که در این زمینه می‌توان انجام داد، به دست آوردن داده‌هایی است که در این پژوهش، امکان گردآوری آن نبوده است. همچنین روش‌های گوناگون دیگری برای شکاف دیجیتال ارائه شده است که می‌توان آن‌ها را نیز آزمایش کرد. از سوی دیگر، روش‌های گوناگون دیگری برای خوشه‌بندی وجود دارد که می‌توان آن‌ها را بررسی نمود. همچنین می‌توان به پارامترهایی پرداخت که بتوانند در جهت کم کردن این شکاف جدی میان استان‌ها مفید باشند.

در این مقاله، به موضوع شکاف دیجیتال و روش‌های بررسی آن پرداخته شد. شکاف دیجیتال، باعث کاهش رشد کلی فناوری اطلاعات می‌شود و همچنین نشان‌دهنده میزان عدالت در عرصه فناوری اطلاعات است. بررسی شکاف دیجیتال درون یک کشور، از موضوع‌های مهم فناوری اطلاعات است که باید به آن دقت نمود. وجود شکاف دیجیتال بسیار زیاد درون یک کشور، نشان‌دهنده همسان نبودن تقسیم امکانات و همچنین رشد نیافتن کلی فناوری اطلاعات در آن کشور است. کاهش شکاف دیجیتال در سطح یک کشور، باعث رشد فناوری اطلاعات درون آن کشور می‌شود.

داده‌های عددی گردآوری شده به روشنی، نشان‌دهنده میزان شکاف دیجیتال در کشور است، ولی خوشه‌بندی، اختلاف ژرف میان استان‌های کشور را به شکل بهتری نشان می‌دهد. با توجه به استان‌های درون هر خوشه، روشن است که امکانات به صورت ناعادلانه‌ای میان استان‌های کشور پخش شده است، به گونه‌ای که تهران همواره در یک خوشه جداگانه قرار می‌گیرد. نیمی از استان‌های کشور در دسته متوسط و نیمی دیگر در خوشه ضعیف قرار می‌گیرند. در پیش آمدن این شکاف بزرگ، هم مردم و هم

مراجع

- اطلاعات، دوره ۲۵، شماره ۲، ۱۳۸۹.
- [۷] شهریاری، حمید، سند راهبردی جامعه اطلاعاتی ایران، تهران: دبیرخانه شورای عالی اطلاع‌رسانی، ۱۳۸۸.
- [۸] رشتیان، سید مسیح‌الله، شهرام مهره‌کش و زهرا عباسی، ارزیابی وضعیت توسعه فناوری اطلاعات در استان‌های کشور با استفاده از شاخص‌های بین‌المللی، دومین کنفرانس بین‌المللی فناوری اطلاعات و دانش، دانشگاه صنعتی امیرکبیر تهران، ۱۳۸۴.
- [۹] مدیر شانه‌چی، محسن، تمرکزگرایی و توسعه‌نیافتگی در ایران معاصر، تهران: مؤسسه خدمات فرهنگی رسا، ۱۳۷۹.
- [۱۰] فتحیان، محمد و معصومه نورزی، ارزیابی آمادگی الکترونیکی مدارس مطالعه‌موردی دبیرستان‌های تهران، سومین کنفرانس بین‌المللی فناوری اطلاعات و دانش، دانشگاه فردوسی مشهد، ۱۳۸۶.
- [1] Ray, joy K., and Acharya, T., *Information Technology: Principles and Applications*, New Delhi: Prentice-Hall of India, 2004.
- [2] Bohn, R. E., and Short, J. E., *How Much Information?: 2009 Report on American Consumers*, San Diego: Global Information Industry Center, University of California, 2009. http://hmi.ucsd.edu/pdf/HMI_2009_ConsumerReport_Dec9_2009.pdf
- [3] Hilbert, M., and López, P., *The World's Technological Capacity to Store, Communicate, and Compute Information*, Science 332(6025): 60-65, 2011.
- [۴] فیدر، جان، جامعه اطلاعاتی، تهران: نشر کتابدار، ۱۳۸۰.
- [۵] مصاحب، لادن، سیستم اطلاع‌رسانی اداری، تهران: سازمان مدیریت صنعتی، ۱۳۸۰.
- [۶] فتاحی، سید رحمت‌الله، جهش علم و فناوری در ایران و جایگاه مراکز اطلاع‌رسانی در پشتیبانی از آن، فصلنامه علوم و فناوری

- [۱۱] حنفی‌زاده، پیام، محمدرضا حنفی‌زاده و سیده ریحانه هدایی‌پور، طراحی مدل ارزیابی آمادگی الکترونیکی دانشگاه‌ها و مؤسسات آموزش عالی ایران، فصلنامه پژوهش و برنامه‌ریزی در آموزش عالی، ۱۴(۲)، ۱۳۸۷.
- [۱۲] تولایی، سید مهبد و فتانه تقی‌یاره، طبقه‌بندی پارک‌های علم و فناوری ایران بر اساس فناوری‌های اطلاعاتی، رشد فناوری ۹، ۱۳۸۵.
- [۱۳] سلیمی‌فرد، خداکرم و مریم عباسی، بررسی آمادگی الکترونیکی شرکت‌های کوچک و متوسط برای تجارت الکترونیکی، پنجمین همایش ملی تجارت الکترونیک، ۱۳۸۷.
- [۱۴] حورعلی، مریم، عباس منتظری، منصوره حورعلی و محمد فتحیان، ارائه مدلی برای سنجش میزان آمادگی الکترونیکی شرکت‌های کوچک و متوسط ایران، نشریه دانشکده فنی دانشگاه تهران، ۱۰۹، ۴۱(۱۰)، ۱۳۸۶.
- [۱۵] سید جوادین، سید رضا، سعید شهباز مرادی، طهمورث حسن قلی‌پور و علی داوری، سنجش آمادگی الکترونیک در معماری منابع انسانی با رویکرد استراتژیک (پژوهشی در شرکت ملی نفت ایران)، نشریه مدیریت فناوری اطلاعات دانشگاه تهران، ۲(۵)، ۱۳۸۹.
- [16] Lievrouw, L. A., and Farb, S. E., *Information and equity. Annual Review of Information Science and Technology*, 37(1): 499-540, 2003.
- [17] Butcher, M., *At the foundations of information justice*, Ethics and Information Technology 11(1): 57-69, 2009.
- [18] Norris, P. .2001. Digital divide: civic engagement, information poverty, and the Internet worldwide. Cambridge; New York: Cambridge University Press.
- [۱۹] قاسمی، میثم، گفت‌وگو با خسرو سلجوقی، معاون فنی سابق شورای عالی اطلاع‌رسانی، IT در چهار سال گذشته تعطیل بود، هفته‌نامه عصر ارتباط، ۷، ۳۲۴، ۱۳۸۸.
- [۲۰] لطفی، میثم، مرگ تدریجی یک رؤیا، ماهنامه تحلیل‌گران عصر اطلاعات، سال چهارم (شماره ۳۱ و ۳۲)، ۱۳۸۹.
- [21] Tien, F. F., and Fu, T.-T., *The correlates of the digital divide and their impact on college student learning*, Computers & Education, 50(1): 421-436, 2008.
- [22] Bruno, G., Esposito, E., Genovese, A., and Gwebu, K. L., *A Critical Analysis of Current Indexes for Digital Divide Measurement*, The Information Society 27(1): 16-28, 2010.
- [23] Emrouznejad, A., Cabanda, E., and Gholami, R., *An alternative measure of the ICT-Opportunity Index*, Information & Management 47(4): 246-254, 2010.
- [24] James, J., *Cumulative bias in the new Digital Opportunity Index: sources and consequences*, Current Science, 92(1): 46-50, 2007.
- [25] Wikipedia, *United Nations Conference on Trade and Development- Wikipedia, the free encyclopedia*, http://en.wikipedia.org/wiki/United_Nations_Conference_on_Trade_and_Development/ (accessed August 10, 2011)
- [26] Partnership on Measuring ICT for Development, *Core ICT Indicators 2010*, Geneva, Switzerland: ITU, UNCTAD and UNESCO, 2010.
- [27] Berthon, P., Pitt, L., Berthon, J.-P., Campbell, C., and Thwaites, D., *e-Relationships for e-Readiness: Culture and corruption in international e-B2B*, Industrial Marketing Management, 37(1), 83-91, 2008.
- [28] International Telecommunication Union (ITU), *Measuring the Information Society, The ICT Development Index 2009 Edition*, Geneva, Switzerland, 2009.
- [29] International Telecommunication Union (ITU), *Measuring the Information Society – The ICT Development Index 2010 Edition*, Geneva, Switzerland, 2010.
- [30] DONG, C., LEI, Z., and LIU, F., *Internet quality abnormal analysis with k-means clustering*, The Journal of China Universities of Posts and Telecommunications 18(2): 94-100, 2011.
- [31] Yiakopoulos, C.T., Gryllias, K.C., and Antoniadis, I.A., *Rolling element bearing fault detection in industrial environments based on a K-means clustering approach*, Expert Systems with Applications 38(3): 2888-2911, 2011.
- [32] Wan Mohd, W.M., Beg, A.H., Herawan, H., and Rabbi, K.F., *An Improved Parameter less Data Clustering Technique based on Maximum Distance of Data and Liloyd k-means Algorithm*, Procedia Technology 1(0): 367-371, 2012.
- [33] Papamichail, G.P., Papamichail, D.P., *The k-means range algorithm for personalized data clustering in e-commerce*, European Journal of Operational Research, 177(3):1400-1408, 2007.
- [34] Jain, A.K., Murty, M.N, Flynn, P.J., *Data clustering: a review*, ACM Comput. Surv., 31(3):264-323, 1999.
- [۳۵] یوسفان، احمد، بررسی تعدادی از عامل‌های تشدیدکننده تمرکزگرایی در فناوری اطلاعات، اولین همایش متخصصان برق و کامپیوتر، دانشگاه شهرکرد، ۱۳۸۹.