

روشی کارا جهت تجمعیع داده ها در شبکه های حسگری سیم با استفاده از آتماتاهای یادگیر

مهندی اثنی عشری	محمد رضا میدبی	محمد احمدی نیا
گروه کامپیوتر، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی	گروه کامپیوتر، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی	گروه کامپیوتر، دانشگاه آزاد
امیرکبیر تهران، ایران	امیرکبیر تهران، ایران	اسلامی، واحد کرمان، ایران
esnaashari@aut.ac.ir	mmeybodi@aut.ac.ir	ahmadinia@gmail.com

حالت هر گره باید بداند که کدامیک از همسایگانش جهت ارسال بسته ها مناسبتر است.

الگوریتمهای مسیریابی زیادی برای شبکه های حسگر ارائه گردیده است. در تعدادی از این الگوریتمها هر گره ممکن است بیش از یک مسیر تا گره سینک داشته باشد که بر اساس یک سری معیارها، یکی از مسیرها انتخاب می گردد. میزان مصرف انرژی در طول مسیر می تواند معیاری مناسب برای این هدف باشد. صرفه جویی در مصرف انرژی به دو شیوه می تواند مدنظر قرار گیرد. یک روش، محاسبه مصرف انرژی برای هر مسیر به صورت جداگانه و سپس انتخاب مسیر با مینیمم انرژی می باشد [1]. مشکل این روش، نیاز به محاسبه انرژی مصرفی برای هر مسیر می باشد. روش دیگر استفاده از تکنیک تجمعیع داده هاست. در تکنیک تجمعیع داده ها، بسته های اطلاعاتی بهم مرتبط، در گره های میانی با هم ترکیب شده و یک بسته را تشکیل می دهند. بدین ترتیب تعداد بسته هایی که در شبکه ارسال می شوند کاهش خواهد یافت و بنابراین انرژی کمتری مصرف خواهد شد.

در این مقاله یک روش تجمعیع داده مبتنی بر آتماتاهای یادگیر ارائه خواهد شد که با شناسایی حسگرهایی که داده های یکسان تولید می کنند و فعال نمودن گره های حسگر به صورت دوره ای از ارسال اطلاعات به فواصل طولانی انرژی زیادی صرف می شود. لذا در بیشتر موارد گره ها از طریق همسایگانشان با گره سینک ارتباط برقرار می نمایند. در این شده است در بخش ۲ خلاصه ای از کارهای انجام شده ارائه

چکیده: یکی از راهکارهای کاهش مصرف انرژی در شبکه های حسگر بی سیم، کاهش تعداد بسته هایی است که در شبکه منتقل می شوند. تکنیک تجمعیع داده ها که داده های به هم مرتبط را با یکدیگر ترکیب می نماید و از ارسال بسته های اضافی در شبکه جلوگیری می نماید می تواند در کاهش تعداد بسته های ارسالی در شبکه موثر باشد. در این مقاله یک روش تجمعیع داده های مبتنی بر آتماتاهای یادگیر ارائه خواهد شد که با شناسایی حسگرهایی که داده های یکسان تولید می کنند و فعال نمودن نود های حسگر به صورت دوره ای، از ارسال بسته های اضافی در شبکه جلوگیری کرده و در مصرف انرژی به میزان زیادی صرفه جویی به عمل می آورد و طول عمر شبکه را افزایش می دهد. نتایج شبیه سازی عملکرد مطلوب روش پیشنهادی را نشان می دهد.

واژه های کلیدی: شبکه های حسگر بی سیم ، طول عمر، تجمعیع داده ها، آتماتاهای یادگیر.

- ۱- مقدمه

شبکه های حسگر از تعداد زیادی حسگر با انرژی محدود تشکیل شده اند. هر گره می تواند اطلاعات خاصی را حس نموده و به همسایگانش ارسال نماید. معمولا یک گره مرکزی که سینک نامیده می شود، مقصد تمام بسته های اطلاعاتی است. برای ارسال اطلاعات به فواصل طولانی انرژی زیادی صرف می شود. لذا در بیشتر موارد گره ها از طریق همسایگانشان با گره سینک ارتباط برقرار می نمایند. در این

جهت استفاده از یادگیری در تجمعی داده ها تلاش‌های اندکی صورت گرفته است. رادی وجک در [7] از یک تکنیک یادگیری ماشین استفاده کرده اند که در آن نودهای حسگر فقط داده های ویژه ای که لازم باشد را به نود سینک ارسال می کنند. الگوریتم یادگیری استفاده شده در این روش، در نود سینک اجرا می گردد و نتایج در شبکه منتشر می شود. لیانگ و همکارانش در [10] برای هر نود از یک Q-learner استفاده کرده اند تا نودها داده هایشان را از مسیری با ماکریزم نرخ تجمعی ارسال نمایند. در [11] روشی برای تجمعی داده ها با استفاده از آtomاتاهای یادگیر ارائه گردیده است.

۳- آtomاتاهای یادگیر

آtomاتای یادگیر ماشینی با حالات محدود است که می تواند تعداد محدودی عمل را انجام دهد. هر عمل انتخاب شده، توسط یک محیط تصادفی ارزیابی شده و پاسخی به آtomاتای یادگیر داده می شود. آtomاتای یادگیر از این پاسخ استفاده نموده و عمل خود را برای مرحله بعد انتخاب می کند. در طی این فرآیند، آtomاتای یادگیر فرا می گیرد که چگونه بهترین عمل را از بین اعمال مجاز خود انتخاب نماید. شکل ۱ ارتباط بین آtomاتای یادگیر و محیط را نشان می دهد.



شکل ۱: ارتباط بین آtomاتای یادگیر و محیط

آtomاتاهای یادگیر با ساختار متغیر: آtomاتای یادگیر با ساختار متغیر توسط ۴ تابی $\{\alpha, \beta, p, T\}$ نشان داده می شود که در آن α مجموعه عمل های آtomات، β مجموعه ورودی های آtomات، p بردار احتمال انتخاب اعمال و $p(n+1) = T[\alpha(n), \beta(n), p(n)]$ الگوریتم یادگیری هستند. در این نوع از آtomاتاهای یادگیر، اگر عمل i در مرحله n ام انجام شود و پاسخ مطلوب از محیط دریافت

می گردد. سپس آtomاتای یادگیر که عنوان استراتژی اصلی یادگیری در الگوریتم پیشنهادی است، در بخش ۳ به اختصار شرح داده می شود. در بخش ۴ الگوریتم پیشنهادی و در بخش ۵ نتایج شبیه سازی ها ارایه می گردد. بخش پایانی مقاله نتیجه گیری می باشد.

۲- کارهای انجام گرفته

در بعضی از کارهای انجام شده در زمینه تجمعی داده ها [2,3] شبکه ابتدا خوش بندی می شود. سپس سرخوشه ها، داده های دریافتی از هر خوش را جداگانه تجمعی می نمایند. لطفی نژاد و لیانگ در [4] سعی کرده اند که تأثیر داده های تا حدی وابسته بروی کارایی روش های خوش بندی در تجمعی داده ها را بررسی نمایند. کسترنین و همکارانش در [5] سعی کرده اند که داده های یک نود در یک بازه زمانی را با یک معادله درجه سوم تخمین زده و به جای ارسال خود داده ها، ضرایب چند جمله ای را ارسال نمایند. داسگوپتا و همکارانش در [6] فرض کرده اند که هرنود در شبکه توانایی تجمعی داده ها را دارد و بر پایه این فرض روشی جهت افزایش طول عمر شبکه ارائه داده اند. بیور و شرف در [7] الگوریتمی را پیشنهاد داده اند که سعی می کند مسیری به یک گروه یکسان از حسگرهای (حسگر هایی که داده های یکسان تولید می کنند) بیاید.

در [3] نودهای حسگر الگوی داده را ارسال می نمایند که نشان می دهد چطور داده های حسگر بر اساس فواصل زمانی از پیش تعیین شده تغییر می کنند. در [4] SPIN با استفاده از مذاکرات فوق داده ای بین نودهای انتقال داده های اضافی حذف می گردد. در [8] انتشار مستقیم به عنوان یک پروتکل تجمعی داده برای شبکه های حسگر پیشنهاد شده است که هدفش نظارت بر رویدادهایی است که معمولاً با استفاده از تعداد کمی از نودها حس می گردد. tag [9] یک سیستم پرس و جوی شبکه های حسگر است که یک سینتکس شبیه sql بکار می برد و از تجمعی داده ها به عنوان یک پرس و جو که در شبکه صورت می گیرد استفاده می کند.

بعضی از قسمتها به دلایلی مثل وزش باد، تابش نور و ... دما موقتاً متفاوت می‌شود.

هدف مسئله این است که حسگرهای موجود در یک ناحیه حسگری یکسان یا به عبارتی اعضای یک ائتلاف یکسان را شناسایی نموده و روشی ارائه دهیم که لازم نباشد تمام حسگرهای موجود در یک ناحیه بخواهند به صورت همزمان اطلاعات را جمع اوری و ارسال نمایند. بلکه در هر زمان یکی از اعضای ائتلاف به عنوان نماینده انتخاب شده و داده های ناحیه را جمع آوری نماید و بقیه گرههای عضو ائتلاف موقتاً غیرفعال گردند.

۴-۲-۴ تشریح روش پیشنهادی

الگوریتم ارائه شده شامل دو فاز می‌باشد که در فاز اول هر گره همسایگان مناسب را با استفاده از تکنیک آtomاتاهای یادگیر شناسایی نموده و ائتلافهای حسگرها مشخص می‌شوند. در فاز دوم، اعضای هر ائتلاف بر اساس یک برنامه زمانبندی به حسگری می‌پردازند.

فاز اول

فاز اول شامل دو مرحله می‌باشد. در مرحله اول هر گره با استفاده از آtomاتاهای یادگیر(LA)، همسایگان مناسب جهت تشکیل ائتلاف را شناسایی می‌نماید. در مرحله دوم با انتشار لیست همسایگان مناسب جهت تشکیل ائتلاف توسط هر گره، ائتلافات کامل تشکیل می‌گردند. جهت شناسایی همسایگان مناسب در هر گره، به ازای هر یک از همسایگانش یک آtomاتای یادگیر در نظر می‌گیریم که دارای دو عمل می‌باشد: گره همسایه متناظر، گرهی مناسبی جهت تشکیل ائتلاف با این نود باشد یا نباشد. در ابتدا هر دو عمل آtomاتا انتقال برابر دارند. در این مرحله هر گره حسگر، داده هایی احتمال برابر دارند. در این مرحله هر گره حسگر، داده هایی را که از محیط حس نموده است، به تمام همسایگانش ارسال می‌نماید. سپس در هر گره، هر یک از آtomاتاهای یادگیر، بر اساس میزان تشابه داده دریافتی از گره متناظر، با داده این گره، پاداش یا جریمه دریافت می‌نماید. جهت پاداش یا جریمه از روابط (۱) و (۲) استفاده می‌گردد.

$$p_i(n+1) = p_i(n) + \alpha[1 - p_i(n)] \quad (1)$$

$$p_j(n+1) = 1 - p_i(n+1)$$

نماید، احتمال $(n)_i$ افزایش یافته و سایر احتمال‌ها کاهش می‌یابند. در هر حال، تغییرات به گونه‌ای صورت می‌گیرد تا حاصل جمع $(n)_i$ ها همواره ثابت و مساوی یک باقی بماند.

۴- تجمیع داده‌ها در شبکه‌های حسگر با استفاده از آtomاتاهای یادگیر

در روشهای تجمیع داده‌ای که تاکنون ارائه گردیده اند، اطلاعات تمام گره‌ها جمع آوری می‌شود و عمل تجمیع داده‌ها در گره‌های میانی انجام می‌گیرد. ولی در بسیاری از کاربردها به دلایلی مثل تشابه محیطی، گره‌هایی که در نزدیکی یکدیگر قرار دارند ممکن است داده‌های یکسانی تولید نمایند. در این حالت می‌توان به جای ارسال داده‌ تمام گره‌ها و سپس تجمیع آنها، گره‌هایی که داده‌های مشابه تولید می‌نمایند را شناسایی نموده و یکی از آنها را فعال نگه داشته و بقیه را غیرفعال نمود تا از ارسال داده‌های افزونه جلوگیری گردد.

۱-۴- بیان مسئله و مفروضات آن

یک شبکه از تعدادی گره حسگر را در نظر بگیرید. فرض کنید که محیط حسگری به ناحیه‌هایی تقسیم بندی گردیده است که داده‌های تولید شده در هر ناحیه مشابهند. لذا حسگرهای موجود در هر ناحیه اطلاعات مشابهی را جمع آوری کرده و ارسال می‌نمایند. در ضمن این نواحی لزوماً ثابت نبوده و در محیط می‌توانند مثل یک ابر یا توده هوا در حرکت باشند، حسگرهایی که در یک ناحیه قرار می‌گیرند، را عضو یک ائتلاف در نظر می‌گیریم. از طرفی فرض کنید که محیط از قبل ناشناخته بوده و لذا حسگرهای عضو یک ائتلاف یکسان یکدیگر را نمی‌شناسند. فرض بر این است که داده‌هایی که توسط گره‌های واقع در یک ناحیه تولید می‌گرددند همیشه و به طور صدرصد یکسان نیستند، بلکه گاه‌آ داده‌هایی که در قسمتهایی از ناحیه تولید می‌گردند، به دلیل وجود نویز، متفاوت از داده‌های تولیدشده در کل ناحیه است. به عنوان مثال یک ابر یا توده هوا را در نظر بگیرید که کل محیط تحت پوشش آن یک دمای خاص دارد. ولی گاه‌ها در

فعال شدن گره‌ها، اعضای ائتلاف بر اساس شماره ترتیب گره‌ها فعال می‌گردند.

از آنجا که فرض شده است نواحی متخرکند ولی گره‌های حسگر متحرک نمی‌باشند، با گذشت زمان ناحیه‌ای که یک حسگر در آن قرار داشته ممکن است تغییر نماید. بنابراین در فواصل زمانی مشخص، الگوریتم مجدد اجرا می‌شود تا ائتلاف‌های جدید دوباره تشکیل گردند.

۵- شبیه سازی

در این بخش کارایی الگوریتم پیشنهادی با آزمایشات مختلف با استفاده از شبیه ساز شبکه J-Sim مورد ارزیابی قرار می‌گیرد. ما عملکرد الگوریتم را با سه روش مختلف مقایسه نموده ایم: ۱) استفاده از روش چندگانمی GPSR [10] و بدون استفاده از تجمعی داده‌ها. ۲) روش ارائه شده در [11] که از Q-Learning استفاده می‌نماید. ۳) روش ارائه شده در [12] که از Learning Automata استفاده می‌کند. معیارهای مد نظر جهت مقایسه عبارتند از: ۱) تعداد کل بسته‌های دریافتی توسط سینک، ۲) کل انرژی مصرفی توسط گره‌ها، ۳) طول عمر شبکه. در آزمایشات، محیط حسگری 60×60 متر و برد رادیویی حسگرها ۱۵ متر و مقادیر پارامترهای a و b برابر با ۰.۱ در نظر گرفته شده است. محیط شامل تعدادی ناحیه می‌باشد که در ابتدا به صورت شکل نشان داده شده در شکل ۲ به ۹ ناحیه 20×20 متر تقسیم گردیده است. داده‌های تولید شده در هر ناحیه مشابهند. نواحی ثابت نبوده و حرکت می‌کنند. بنابراین شکل محیط تغییر می‌کند. جهت شبیه سازی این موضوع ما با فواصل زمانی ۲۵۰ دقیقه، شکل نواحی محیط را تغییر می‌دهیم. در ضمن به دلیل ماهیت پویای محیط، باید الگوریتم تشکیل ائتلاف با فواصل زمانی مشخص اجرا گردد (در اینجا ۱۳۰ دقیقه).

$$p_i(n+1) = (1-b)p_i(n) \quad (2)$$

$$p_j(n+1) = 1 - p_i(n-1)$$

در هر دور، هر گره پس از انتخاب یک عمل به ازای هر آutomata و دریافت داده‌های گره‌های مجاور، بر اساس رابطه (۳) به ازای هر آtomata یادگیر سیگنال تقویتی β را محاسبه می‌نماید. اگر β برابر با یک شود عمل انتخابی با استفاده از رابطه (۱) جریمه می‌گردد. و اگر برابر با صفر شود طبق رابطه (۲) پاداش دریافت می‌کند.

$$\beta_i(n) = \begin{cases} 0 & : |D_{rec}^i(n) - D_{sens}(n)| < thresh \\ 1 & : otherwise \end{cases} \quad (3)$$

به طوری که $D_{rec}^i(n)$ داده دریافتی از همسایه i در دور n و $D_{sens}(n)$ داده حس شده توسط این گره در دور n می‌باشد. $thresh$ یک مقدار استانداری است.

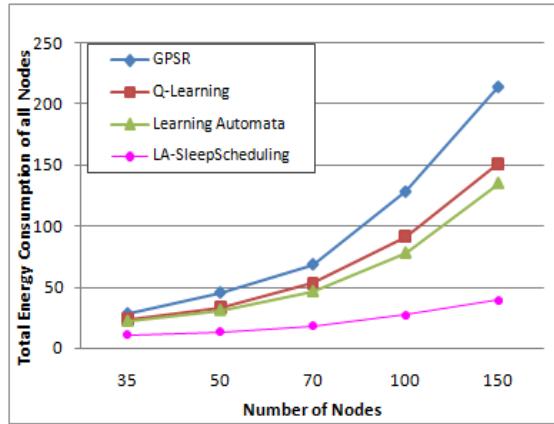
این مرحله در چندین دور انجام می‌گیرد. در انتها هر گره آن دسته از همسایگانش را که جهت تشکیل ائتلاف مناسب می‌باشد را شناسایی نموده است.

در مرحله دوم از این فاز هر گره لیست همسایگان مناسب جهت تشکیل ائتلاف را به آن همسایگان ارسال می‌نماید و هر گره که لیست گره‌های هم ائتلافی را دریافت نموده با لیست زیر ائتلاف خودش ادغام و لیست جدید را به تمام همسایگان هم ائتلافی ارسال می‌نماید. این مرحله در چندین دور انجام گرفته و در انتها هر گره لیست کامل گره‌های هم ائتلافی خود را دارد.

فاز دوم

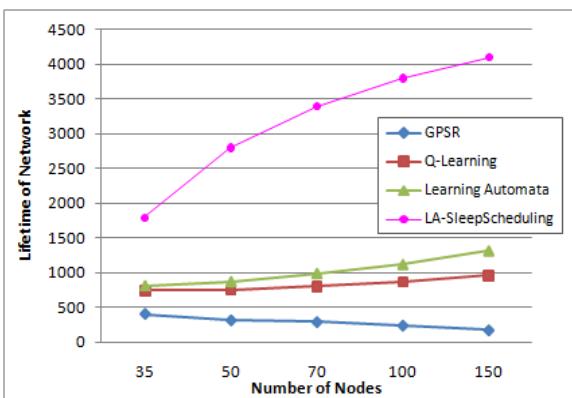
فاز دوم، فاز حالت پایدار شیوه حسگر است که در این فاز، گره‌ها داده‌های محیط را جمع آوری کرده و به گره سینک ارسال می‌نمایند. ولی لزومی ندارد که تمام گره‌ها فعال باشند، بلکه کافی است که در هر ناحیه فقط یک گره فعال بوده و داده‌ها را جمع آوری و ارسال نماید. در این فاز از ائتلافاتی که در فاز اول تشکیل گردیده استفاده می‌گردد. در هر ائتلاف یک برنامه زمانبندی ایجاد می‌گردد تا اعضای ائتلاف بر اساس این برنامه زمانبندی به نوبت فعال شده و داده‌ها را ارسال کنند. جهت جلوگیری از تداخل در زمان

میزان انرژی مصرفی در شبکه در روش پیشنهادی کمتر از روش‌های دیگر می‌باشد.



شکل ۴: مقایسه کل انرژی مصرفی توسط نودها در روش‌های مختلف

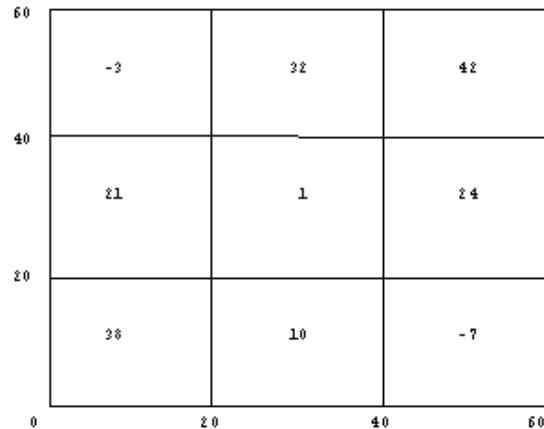
در آزمایش سوم طول عمر شبکه در استفاده از روش‌های مختلف تجمعیع مورد مقایسه قرار گرفته است. نتیجه ارزیابی در شکل ۵ نشان داده شده است که نشان دهنده عملکرد بسیار مطلوب روش پیشنهادی نسبت به دیگر روش‌های تجمعیع می‌باشد.



شکل ۵: مقایسه طول عمر شبکه در روش‌های مختلف تجمعیع

۶- جمع بندی

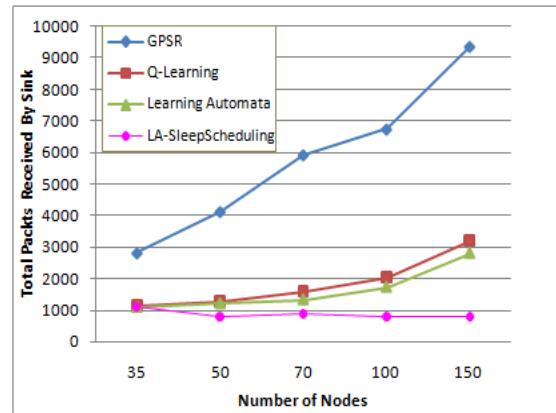
در این مقاله روشی جدید مبتنی بر آtomاتای یادگیر، جهت تجمعیع داده‌ها در شبکه‌های حسگری سیم معرفی گردید. در این روش با غیر فعال کردن گره‌های غیرضروری در شبکه، در مصرف انرژی شبکه به میزان زیادی صرفه جویی به عمل آمده و طول عمر شبکه افزایش می‌یابد. نتایج



شکل ۲: تقسیم بندی محیط حسگری به ۹ ناحیه

در آزمایش‌های اول و دوم کل زمان شبیه سازی ۱۰۰۰ دقیقه در نظر گرفته شده و در آزمایش سوم شبیه سازی تا از بین رفتن اولین گره ادامه می‌یابد. احتمال وجود نویز در محیط، ۵ درصد در نظر گرفته شده است. آزمایشات برای تعداد گره‌های حسگر برابر با ۱۵۰, ۱۰۰, ۷۰, ۵۰, ۳۰ ۱۵۰, ۱۰۰, ۷۰, ۵۰, ۳۰ انجام گرفته است و هر آزمایش ۱۰ بار تکرار گردیده است. نیز برابر با ۲۰ می‌باشد.

در ازمایش اول تعداد بسته‌هایی که توسط گره سینک دریافت می‌گردند در روش‌های مختلف با یکدیگر مقایسه شده‌اند (شکل ۳). روش پیشنهادی عملکردی به مرتبه نیز نسبت به دیگر روش‌های تجمعیع دارد.



شکل ۳: مقایسه تعداد کل بسته‌های دریافتی توسط گره سینک

در آزمایش دوم کل انرژی که توسط گره‌ها در مدت شبیه سازی مصرف شده است، مورد مقایسه قرار گرفته است که نتایج در شکل ۴ نشان داده شده است. با توجه به شکل،

آزمایشات نشان دهنده عملکرد مطلوب روش ارائه شده می باشد.

مراجع

- Obradovic, Learning from Class-Imbalanced Data in Wireless Sensor Networks, IEEE Semiannual Vehicular Technology Conference, VTC-Fall 2003, Vol. 5, pp. 3030-3034, Orlando, Florida, U.S.A., October 2003.
- [8] Y. Xu, W. C. Lee, J. Xu, and G. Mitchell, Processing Window Queries in Wireless Sensor Networks, IEEE International Conference on Data Engineering (ICDE'06), Atlanta, GA, April 2006.
- [9] R. Rosemark and W. C. Lee, Decentralizing Query Processing in Sensor Networks, the Second International Conference on Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous'05), San Diego, CA, July, 2005, pp. 270-280.
- [10] B. Karp and H. Kung, Greedy perimeter stateless routing for wireless networks, In Proceedings of the Sixth Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom 2000), pages 243–254, Boston, MA, August 2000.
- [11] P. Beyens, M. Peeters, K. Steenhaut and A. Nowe, *Routing with Compression in Wireless Sensor Networks: a Q-learning Approach*, In "Fifth European Workshop on Adaptive Agents and Multi-Agent Systems (AAMAS 05), Paris, France.", 2005.
- [12] Y. Xu, W. C. Lee, J. Xu, and G. Mitchell, Processing Window Queries in Wireless Sensor Networks, IEEE International Conference on Data Engineering (ICDE'06), Atlanta, GA, April 2006.
- [1] R. Shah and J. Rabaey, Energy Aware Routing for Low Energy Ad Hoc Sensor Networks, in Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC), Orlando, Florida, March 2002.
- [2] R. Virrankoski and A. Savvides, TASC : Topology Adaptive Spatial Clustering for Sensor Networks, Second IEEE Intl. Conf. on Mobile Ad Hoc and Sensor systems", Washington, DC, November, 2005.
- [3] S. Soro and W. Heinzelman, Prolonging the Lifetime of Wireless Sensor Networks via Unequal Clustering, Proceedings of the 5th International Workshop on Algorithms for Wireless, Mobile, Ad Hoc and Sensor Networks (IEEE WMAN '05), April 2005.
- [4] M. Lotfinezhad and B. Liang, Effect of partially correlated data on clustering in wireless sensor networks, in Proceedings of the IEEE International Conference on Sensor and Ad hoc Communications and Networks (SECON), Santa Clara, California, 2004..
- [5] C. Guestrin, P. Bodik, R. Thibaux, M. Paskin and S. Madden, Distributed Regression : An Efficient Framework for Modeling Sensor Network Data, 2004.
- [6] K. Dasgupta, K. Kalpakis and P. Namjoshi, An Efficient Clustering-based Heuristic for Data Gathering and Aggregation in Sensor Networks, IEEE Wireless Communications Conference, Vol.4, No. 1, 2003.
- [7] P. Radivojac, U. Korad, K. M. Sivalingam and Z.