

ИЗСЛЕДВАНЕ НА АНОМАЛИИТЕ В МЕСЕЧНАТА КОНСУМАЦИЯ НА ЕЛЕКТРИЧЕСТВО В ЗАВИСИМОСТ ОТ ТЕМПЕРАТУРАТА¹

Андрей Бъчваров¹, Калоян Наралампиев², Петко Русков³

¹Balkans Investment Consulting Agency Ltd., 83-85 James Baucher Blvd, floor 2, office 7
1407 Sofia, Bulgaria, andrey.bachvarov@bica-bg.com

²Department of Software Engineering, Sofia University, Bulgaria
125 Tzaridkoshose Blvd., bl.4,1113 Sofia, Bulgaria; k.haralampiev@hotmail.com

³Department of Software Engineering, Sofia University, Bulgaria
125 Tzaridkoshose Blvd., bl.2,1113 Sofia, Bulgaria; petkor@fmi.uni-sofia.bg

Abstract: In developing countries like Bulgaria, currently electricity consumption climbed steadily, mainly due to increased access to electrical appliances, dependence on electrical power for climate control, introduction of expensive renewable energy sources and the same time artificially low prices and related. Governments and electricity power providers find it increasingly difficult to respond to the “new challenges” by developing new consumption model but due to the overwhelming complexity of data, most of those models are basically at the macro level and it is almost impossible to identify individual consumption patterns that could generate useful business process innovation. Households decide intuitively on their consumption due to lack of access to necessary tools or guidance to plan accordingly. Our research demonstrates that hidden patterns exist in the clients’ consumption data depending on specific parameters of their respective environment and temperature. Our approach relies on data processing of large scale sample data base provided by the “CEZ Electro Bulgaria” AD. and data mining using IBM SPSS Modeler predictive models and tools. This paper identifies and highlights typical consumption patterns within a database of randomly selected clients provided by CEZ. The results include identification and visualization of electricity consumption patterns based on Anomaly node method. The generated patterns recognition and predictions are useful for researchers as they can use those findings for informed decision making and argued business process innovation versus “blind” and “intuitive” actions. Electric distribution companies can further use the results specific business transformations and process innovations including planning of delivery, network optimization and maintenance, proper cost allocation to most problematic areas, demand forecast based on improved supply predictions and related.

Резюме: Потреблението на ел. енергия расте непрекъснато през последните години. Правителствата и доставчиците на електроенергия търсят решения на това предизвикателство чрез промени в регулациите, поощряване на иновациите и новите бизнес модели, както за производителите, така и при потребителите. Повечето от инициативите са на макро ниво, като домакинствата обикновено нямат необходимите компетенции и инструменти за енергийно ефективни решения. Представеното изследване е опит да се покаже, че съществуващите добри практики и технологии позволяват да се откриват закономерности и тенденции, като по този начин се подпомогне повишаването на енергийната ефективност. Представения модел за изследване на аномалиите в месечната консумация на електричество в зависимост от температурата се базира на обработката на голям обем данни, представени от ЧЕЗ Електро България АД и средните стойности на температурата в съответните населени места. Използван е програмния продукт на фирмата IBM SPSS Modeler.

Key words: electricity consumption, anomaly analysis, predictive models, process innovation, energy efficiency

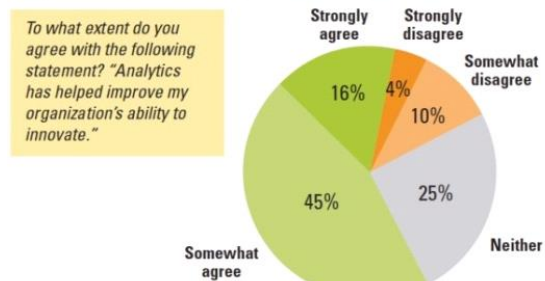
Ключови думи: потребление на ел. енергия, аномална консумация, иновации на процеси, енергийна ефективност

ВЪВЕДЕНИЕ

Потреблението на ел. енергия в световен и местен план расте непрекъснато през последните години [4, 10]. Правителствата и доставчиците на електроенергия търсят решения на това предизвикателство чрез промени в регулациите, поощряване на иновациите и новите бизнес модели, както за производителите, така и при потребителите [3, 5, 19, 22]. Повечето от инициативите са на макро ниво, като домакинствата обикновено нямат необходимите компетенции и инструменти за енергийно ефективни решения [1,2]. Идеята, че събраните данни от организациите и аналитичността могат да се използват за постигане на конкурентно предимство и технологични иновации ¹е базова за водещите фирми – фиг.1.[15, 16, 18].

ANALYTICS FACILITATES INNOVATION

Sixty-one percent of our respondents somewhat or strongly agree that analytics has improved their organizations’ ability to innovate.



Фиг. 1 Отговори на въпроса „Помага ли аналитичността да се подобри възможностите на организацията за иновации?“

¹ Този доклад съдържа предварителни резултати от проект “Методи и модели за подобряване на ефективността при хибридни електроснабдителни системи от възобновяеми енергийни източници с малка мощност” представен за „Финансиране на научни и научноприложни изследвания в приоритетните области”, Вх. Номер FFNNIPO_12_00914

Традиционно иновацията се разбира като нещо ново, творческо, като възможността да генерират идеи, но на практика иновацията може да се управлява, така както се управлява например качеството [8]. Повечето от организациите не страдат от липса на нови идеи, а от липсата на компетенции да управляват процеса на

използване на иновацията и комерсиализирането ѝ. Решението много често е просто и се състои в създаване на творческа среда за управление на процеса, който включва хората, технологиите и финансите. Използването на съвременните научни и технологични постижения позволява на организациите проактивно да управляват цикъла за иновации и комерсиализиране на откритията и технологиите.

Водещите фирми използват иновативни бизнес модели в допълнение към иновациите на продуктите и процесите си [6, 8, 20, 23]. Иновациите на бизнес моделите се развива в следните направления:

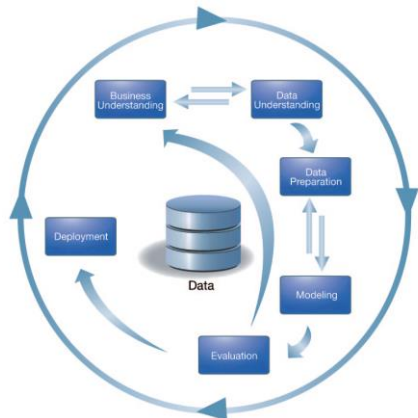
- Добавяне на иновативни дейности в бизнес процесите си;
- Съвързване на дейностите по иновативен начин;
- Иновативно промяна на доставчици или партньори

Целта на представеното изследване е да се покаже, че съществуващите добри практики и технологии позволяват да се откриват закономерности и тенденции, чрез извличане на знания от събраните данни, като по този начин се подпомогне вземането на решение, а от там и повишаването на енергийната ефективност. То е продължение на представеното от авторите в [7] като аномалиите в консумацията се анализират и в зависимост от месечните температури в съответните населени места. Представения модел за изследване на аномалиите в месечната консумация на електричество в зависимост от температурата се базира на обработката на голям обем данни, представени от ЧЕЗ Електро България АД и средните стойности на температурата в съответните населени места [11].

МЕТОДОЛОГИЯ НА ИЗСЛЕДВАНЕТО

Изследването се основава на CRISP-DM модела за извличане на информация [13,14, 21]. Чрез обработване на събраните от организацията голям обем от данни и използване на програмния продукт на фирмата IBM SPSS Modeler се изграждат модели, които позволяват да се анализират данните и се открият нови успешни образци (patterns) за постигане на бизнес целите на организацията. Методологията предлага софтуерни инструменти за обработване на информацията и моделиране на следните етапи – фигура 2:

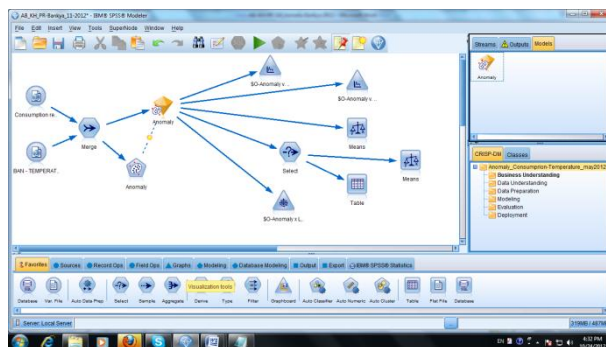
1. Познание на бизнеса
2. Оценка на данните
3. Подготовка на данните
4. Моделиране
5. Оценка
6. Внедряване



Фиг. 2. Етапи на CRISP-DM методологията [13]

За целите на изследването е проектиран „поток” (stream) на обработване на данните – фиг. 3. Той включва:

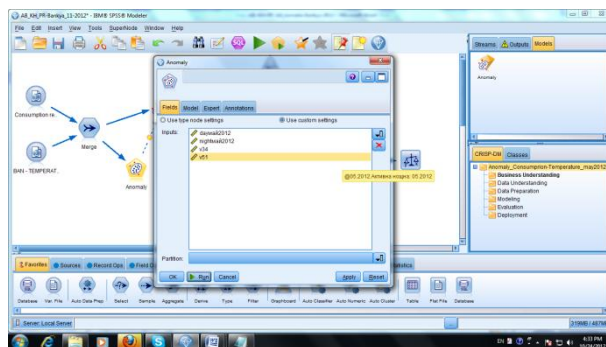
- Зареждане на данните от базите за консумация на ел. енергия и средномесечни температури на населените места;
- Обединяване на базите от данни;
- Използване на метода за откриване на аномалии в консумацията на енергия в зависимост от температурите;
- Визуализиране на резултатите чрез таблици и графики.



Фиг. 3. „Поток” на обработване на данните

Аномалиите в консумацията на отделни потребители на ел. енергия се открива чрез анализ на месечното потребление за последните 2 години и средномесечните температури за съответните населени места.

След провеждане на серии от експерименти с проектирания и представен на фиг. 3 „поток” бяха открити потребителите с аномални стойности на консумираната енергия. В този доклад се представят само част от резултатите – за консумация на ел. енергия по дневна и нощна тарифа през месец май 2012, който е последният месец, за който имаме налични данни – фиг.4.



Фиг. 4. Входни данни за откриване на аномалиите

ОБСЪЖДАНЕ НА РЕЗУЛТАТИТЕ

Откриването на аномалии е следващата стъпка след клъстерния анализ [7]. За аномални се считат всички потребители, които не попадат в нито един от установените клъстери, като процедурата на самия клъстерен анализ е „скрита” и като краен резултат се вижда само дали съответният потребител има „нормално” или аномално потребление.

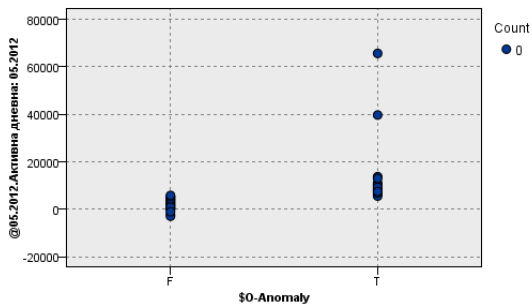
За да се опишат двете групи потребители, в „потока” се добавя възел (node) за сравняването на средните им стойности – фиг. 5.

Field	F*	T*	F-Test	df	Importance
@05.2012.Активна дневна: 05.2012	504.066 818.099 20.834 1542	13088.524 13931.839 3040.176 21	1042.423	1, 1561	1,000 Important
@05.2012.Активна нощна: 05.2012	165.521 297.809 7.584 1542	5742.857 8543.533 1864.352 21	630.129	1, 1561	1,000 Important

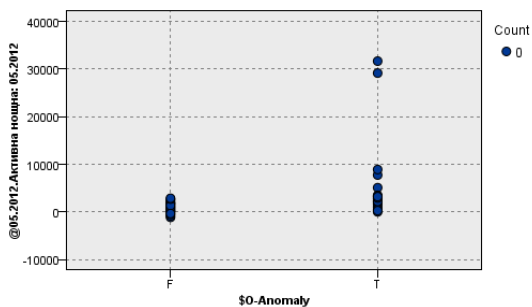
Фиг. 5. Описателни характеристики на групите потребители

На фиг. 5 е показана средната консумация на ел. енергия през месец май 2012 на двете групи потребители – тези с „нормално” потребление (F – False) и тези с аномално потребление (T – True). Видно е, че аномалните потребители са само малка част от всичките – 21 човека. Тяхната средна дневна консумация през месец май 2012 обаче е била 26 пъти по-висока от консумацията на „нормалните” потребители, а средната им нощна консумация е била 35 по-голяма отколкото на „нормалните”. В същото време групата на аномалните потребители е доста по-нехомогенна от групата на „нормалните”. Разликата в стандартните отклонения между двете групи е 17 пъти при дневната консумация и 29 пъти при нощната консумация.

Разликите в средните стойности и в разсейването могат да се представят и графично, като от големия и разнообразен набор от възможности на IBM SPSS Modeler за визуализиране на данните сме избрали възел за графики и сме го добавили в „потока”. Като резултат се получават две графики за всяка от двете тарифи – фиг. 6 и фиг. 7.



Фиг. 6. Дневна консумация през месец май 2012 на двете групи потребители



Фиг. 7. Нощна консумация през месец май 2012 на двете групи потребители

Фиг. 6 и 7 отново потвърждават многократно по-високите средни стойности и многократно по-големите разсейвания в групата на аномалните потребители.

Освен като група със съответната средна консумация и съответното разсейване, аномалните потребители са важни и като индивидуални случаи. Това е така, защото те се отличават от общата маса потребители и биха могли да бъдат обект на проверка от страна на ЧЕЗ Електро, за да се установи дали тяхната аномална консумация е в резултат на обективни причини или става въпрос за злоупотреби и/или кражба на електричество.

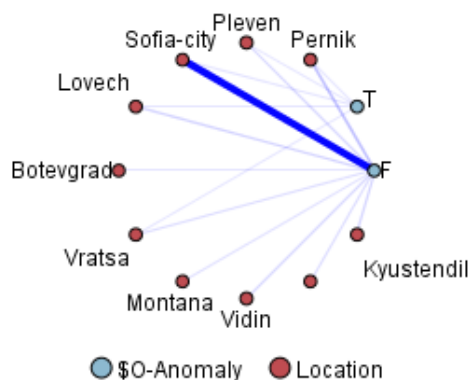
За да идентифицираме аномалните потребители към „потока” добавяме възел за селектиране, последван от възел за таблица. Полученият резултат е представен на фиг. 8.

Location	dayjanu2	dayfevru2	daymart2011	dayapril20	dayмай2011	dayюни2011	dayюли2011	id
1 Lovech	0.5	1.5	7.8	12.5	18.0	22.2	25.2	
2 Pernik	0.2	1.3	6.2	10.5	15.5	19.7	22.5	
3 Pernik	0.2	1.3	6.2	10.5	15.5	19.7	22.5	
4 Pernik	0.2	1.3	6.2	10.5	15.5	19.7	22.5	
5 Plevan	-0.5	1.0	7.5	12.9	18.2	22.4	25.0	
6 Plevan	-0.5	1.0	7.5	12.9	18.2	22.4	25.0	
7 Sofia-city	-1.0	0.8	6.4	11.2	16.2	20.2	23.5	
8 Sofia-city	-1.0	0.8	6.4	11.2	16.2	20.2	23.5	
9 Sofia-city	-1.0	0.8	6.4	11.2	16.2	20.2	23.5	
10 Sofia-city	-1.0	0.8	6.4	11.2	16.2	20.2	23.5	
11 Sofia-city	-1.0	0.8	6.4	11.2	16.2	20.2	23.5	
12 Sofia-city	-1.0	0.8	6.4	11.2	16.2	20.2	23.5	
13 Sofia-city	-1.0	0.8	6.4	11.2	16.2	20.2	23.5	
14 Sofia-city	-1.0	0.8	6.4	11.2	16.2	20.2	23.5	
15 Sofia-city	-1.0	0.8	6.4	11.2	16.2	20.2	23.5	
16 Sofia-city	-1.0	0.8	6.4	11.2	16.2	20.2	23.5	
17 Sofia-city	-1.0	0.8	6.4	11.2	16.2	20.2	23.5	
18 Sofia-city	-1.0	0.8	6.4	11.2	16.2	20.2	23.5	
19 Sofia-city	-1.0	0.8	6.4	11.2	16.2	20.2	23.5	
20 Vratsa	0.4	1.2	7.1	12.7	17.4	21.9	24.6	
21 Vratsa	0.4	1.2	7.1	12.7	17.4	21.9	24.6	

Фиг. 8. Изходни данни на аномалните потребители

В таблицата на фиг. 8 са останали само тези 21 потребителя, които имат аномална консумация. Те присъстват в таблицата с пълната информация за себе си – населено място, средномесечни температури, месечна консумация, клиентски номер и т.н. Това дава възможност, първо, да се идентифицират аномалните потребители, и второ, да се види какво точно ги отличава от останалите и ги прави аномални.

След като вече са идентифицирани аномалните потребители те могат да се изследват по множество други признаци. Например, на фиг. 9 сме обвързали „нормалното” и аномалното потребление с населеното място.



Фиг. 9. Връзка между „нормалното” и аномалното потребление и населеното място

От фиг. 9 става ясно, че аномални потребители има в пет от населените места – Перник, Плевен, София, Ловеч и Враца. В Ботевград, Монтана, Видин, Благоевград и Кюстендил всички потребители са „нормални”.

След като установихме, че аномалните потребители се позиционират в пет населени места, можем да видим дали тяхната консумация се различава в зависимост от населеното място. За целта отново добавяме възел за сравняването на средните стойности към „потока”.

Field	Lovetch*	Pernik*	Pleven*	Sofia-city*	Vratsa*	F-Test	df	Importance
@05.2012 Активна дневна: 05.2012	8845.000	6675.000 1366.700 783.291	12087.000 2248.771 1588.000	16076.385 17136.160 4752.713	6511.000 1301.076 920.000	0.394	4, 16	0.190 Unimportant
@05.2012 Активна нощна: 05.2012	5062.000	1791.667 898.488 518.742	8273.000 851.357 602.000	6756.615 10675.560 2960.868	2890.500 392.444 277.500	0.262	4, 16	0.102 Unimportant

Фиг. 10. Описателни характеристики на аномалните потребители по населени места.

Резултатите на фиг. 10 показват, че аномалните потребители от Плевен (2) и София (13) имат най-висока консумация (както дневна, така и нощна). На следващо място е аномалният потребител от Ловеч, а с най-ниска консумация са аномалните потребители от Перник (3) и Враца (2).

ЗАКЛЮЧЕНИЕ И БЪДЕЩИ ИЗСЛЕДВАНИЯ

Представените резултати включват анализ на аномалиите при месечно потребление на ел. енергия с отчитане на средномесечните температури. Методиката позволява откриване на конкретни потребители с аномалии на потреблението при съответната тарифа и вземане на решения за повишаване на енергийната ефективност. Получените модели и резултати дават възможност да се направят иновативни промени в бизнес процесите на организацията. Планирани са извършване на следните дейности:

- допълнителни изследвания и социологически проучвания за демографските характеристики на потребителите на ел. енергия, специфични икономически и социални характеристики и други параметри, които влияят на потреблението;
- разширяване на анализите за аномалии и с демографски и други социални и икономически данни;
- реални проверки на потребителите с показани аномалии в потреблението и анализ на процесите;
- моделиране на иновативните бизнес модели и процесите, свързани с тях;
- генериране на идеи, разработване на прототипи и доказване на възможностите за комерсиализация на тези идеи;
- иновация на бизнес процесите по цялата верига на потребление, отчитане, заплащане и контролиране.

REFERENCES

1. Бъчваров, А., Петко Русков, Енергийна ефективност – състояние, потребителско поведение и предизвикателства, *Automatics and Informatics* '11, 3-7.10.2011, ISSN 1313-1850, стр. 21-28.
2. Втори национален план за действие по енергийна ефективност 2011 – 2013 г
<http://www.government.bg/cgi-bin/cms/vis/vis.pl?s=001&p=0228&n=2831&g=>
3. Русков П., Танев С., Откриване на иновативни шаблони, Международна научна конференция, приложна информатика и статистика – съвременни подходи и методи, Равда, 25-26 септември 2009.
4. A-L Lindén and A Carlsson-Kanyama, Energy Efficiency in Residences - A challenge for women and men, *Energy Policy*, 35, 2007, pp. 2163-2172.

5. Ambuj D. Sagara, Bob van der Zwaana, Technological innovation in the energy sector: R&D, deployment, and learning-by-doing, *Energy Policy* 34 (2006) 2601–2608.
6. Amit R., Christoph Zott, Creating Value Through Business Model Innovation, *MIT SLOAN MANAGEMENT REVIEW*, Spring 2012, Reprint 53310, Vol. 53, No 3, pages 41-49.
7. Bachvarov A., Petko Ruskov, Kaloyan Haralampiev, Electricity end users' consumption patterns in different west Bulgarian locations, *UNITECH'12 International scientific conference*, 16 – 17 November 2012, Gabrovo.
8. Baldwin E., Innovation Isn't 'Creativity,' It's a Discipline You Manage, *MIT SLOAN MANAGEMENT REVIEW*, April 2010.
9. Bojda, Nicholas, Business Case for Energy Efficiency in Support of Climate Change Mitigation, Economic and Societal Benefits in the United States, *Lawrence Berkeley National Laboratory*, 09-07-2011, <http://escholarship.org/uc/item/7xm4m8ht>
10. Carl Haigney, Customer Relationships and the Economy, Shaping a New Era in Energy, *www.UtilitiesProject.com*.
11. CEZ Electro Bulgaria AD
<http://www.cez.bg/en/home.html>,
<http://www.cez.bg/en/customer-service/ways-of-payment.html>.
12. Druckman A., T. Jackson, Household energy consumption in the UK: A highly geographically and socio-economically disaggregated model, *Energy Policy*, - Elsevier, 2008.
13. IBM Corp., CRISP-DM 1.0, Step-by-step data mining guide, © Copyright IBM Corporation 2010.
14. IBM Corp., Switching perspectives, creating new business models for a changing world of energy, 2010.
15. Kiron D., Pamela Kirk Prentice, Renee Boucher Ferguson, *Innovating With Analytics*, *MIT SLOAN MANAGEMENT REVIEW*, VOL.54 NO. 1, 2012
16. LaValle S., Michael Hopkins, Eric Lesser, Rebecca Shockley and Nina Kruschwitz, *Analytics: The new path to value*, How the smartest organizations are embedding analytics to transform insights into action, *IBM Institute for Business Value*, *MIT Sloan Management Review*, 2010.
17. Letschert V., The Boom of Electricity Demand in the Residential Sector in the Developing World and the Potential for Energy Efficiency, *Lawrence Berkeley National Laboratory*, 02-05-2010.
18. Lukas G. Swan, V. Ismet Ugursal, Modeling of end-use energy consumption in the residential sector: A review of modeling techniques, *Renewable and Sustainable Energy Reviews* 13 (2009) 1819–1835, 2009.
19. Marc Teerlink and Michael Haydock, Customer analytics pay off, *Driving top-line growth by bringing science to the art of marketing*, Copyright IBM Corporation 2011.
20. Miles I., Patterns of innovation in service industries, *IBM Systems journal*, vol.47, No 1, 2008, pp.115-128.
21. Morelli T., Shearer C., Buecker A., *IBM SPSS predictive analytics: Optimizing decisions at the point of impact*, *IBM redpaper 4710*, Copyright IBM Corp. 2010.
22. Pearson N., Eric Lesser and Joel Sapp, *IBM Global Business Services, Executive Report*, *IBM Institute for Business Value*, A new way of working, Insights from global leaders, 2010.
23. Ruskov P., A. Bachvarov, Disruptive innovations for the energy system, *Anniversary Scientific Conference with International Participation "40 Years Department of Industrial Automation"*, 18 March 2011, Sofia, Bulgaria, pp. 141-144.