©2017 Published in 5th International Symposium on Innovative Technologies in Engineering and Science 29-30 September 2017 (ISITES2017 Baku - Azerbaijan)

On The Development of The Autonomous City

[1]Yusuf Mücahit Çetinkaya, [2]Zahit Saygın Doğu and [2]Orçun Dayıbaş
[1]Department of Computer Engineering, Middle East Technical University, Ankara, Turkey
[2] HAVELSAN Air Electronic Industry and Trade Inc., METU Teknokent, Ankara, Turkey

Abstract

Human environments like cities are traditionally accepted as static systems. This perception makes the transition to smart cities harder due to sustainability issues. As a phenomenon, "smart city" defines a dynamic environment. Furthermore, emerging technologies affect the management approach of the cities like many other systems. These technologies change the attitude of the stakeholders in independent ways that are not necessarily compatible. The key point on value adding smart systems of a smart city is governing all these systems in coordination. Therefore, operation of these systems allows a smart city to be transformed into an autonomous entity. In this paper, we discuss the requirements and challenges of developing a smart city as a self-learning ICT system. Our solution approach is also introduced in the paper to develop an ultimate smart city namely "the autonomous city".

Key words: Smart city, Autonomous City, Self-learning, Integration, Orchestration

1. Introduction

Cities are one of the most important subjects in population studies. Recent studies show that the urban population is increasing over time [1]. The rapid growth of the urban population gives rise to many problems in the governance of the cities. On one hand, city officials are in search of the answers to these emerging problems and policy makers also seek new solutions in terms of regulations. On the other hand, the technology is rapidly being developed throughout the years. Therefore, the policy makers tend to adopt these improvements and integrate the resulting mechanisms to the city infrastructure in order to overcome the aforementioned problems.

City infrastructure consists of multiple interconnected complex systems and sub-systems. Therefore, in order to design a system to work with the infrastructure of the city, excessive thought on integration must be given. If a system is integrated with a newly developed instrument, other systems can be stale in terms of the technology. The worse is that different systems can be integrated with a variety of infrastructural systems and when they need to co-operate, the whole system may fail eventually.

There are different domains on managing the city which requires varying components to be integrated. These domains can be summarized as energy grids, public lighting, natural resources, water management, waste management, environment, transport, mobility, logistics, office and residential buildings, healthcare, public security and so on [2]. Just by connecting components together, it is not safe to assume that the integrated subsystems will work together seamlessly.

*Corresponding author: Address: Department of Computer Engineering, Middle East Technical University, Ankara, Turkey

2. Preliminaries

2.1. Reinforcement learning

Machine learning can be divided into categories like clustering, classification and regression problems etc. Reinforcement learning is one of these categories [3]. In reinforcement learning there is an agent and an environment. The agent can observe the environment and knows its state and the possible actions it can take, given that state. When the agent interacts with the environment, the agent uses both environment and its current state as input. Committing an action results in a positive or negative reward value associated with it. In order to increase the reward according to this input, it chooses the best possible action that it knows. To do so, the agent explores the possible states and then have an idea of what is a better action to take after learning the consequences by trial. More specifically, the reward values are calculated by summing up the instant and future reward values that the agent gets after taking that action.

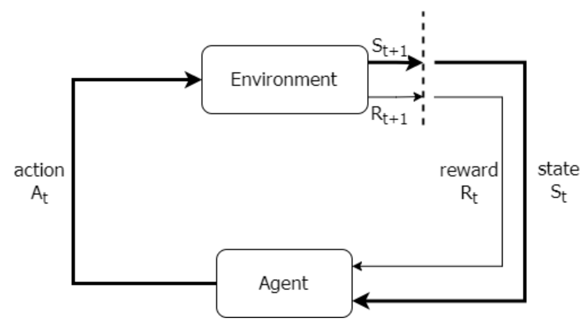

Figure 1. The reinforcement learning flow. Agent takes the current state and a reward from the environment as input. According to this input, it decides on an action that it knows already and waits for the next state and reward.

Many reinforcement learning problems can be reduced to k-armed bandit problem [4]. This problem can be summarized as the agent inside a room with ***k*** gambling machines. Each machine gives a reward according to distribution parameters inside it. The aim of the agent is to get more rewards from these machines. However, the agent has permission of only ***h*** number of pulls. This restriction forces the agent to make the actions more wisely. This problem shows the significance of the tradeoff between the exploration and the exploitation. It is simply decision between choosing the action that the agent knows best, which is called exploitation, and trying others to learn them, namely exploration. This problem should be tackled according to the length of the game. The longer the game lasts the worse the consequences of prematurely converging on a sub-optimal arm, so the agent should explore more [3].

During the city management there is no opportunity to have many changes, so the 'game' in the context of the cities are long term, and the nature does not allow to explore much. This dilemma, therefore, is a major concern while forming an autonomous smart city. Therefore, virtually exploring the possible outcomes enables the system to optimize the decisions.

2.2. Simulation technologies

Machine learning applications needs data in order to recognize the patterns in the data and give successful results on the tasks such as classifications, regression values or actions. In the context of smart cities, there is inherently huge amount of data. It is for the very reason that makes it suitable for machine learning. However, in order to enable an effective learning system, the data must be evenly distributed and unbiased. [5]. Data that is coming from the ICT infrastructure of the city is biased [6]. That is to say, the data that comes from the sensors are generated from the real environment and they consist of limited number of instances of cause and effect pairs. Since the real life actions cannot be repeated or changed, the data will be biased. To address this issue, a realistic simulator can be used for repeating and changing the actions that have been taken in order to see the differences in the effects.

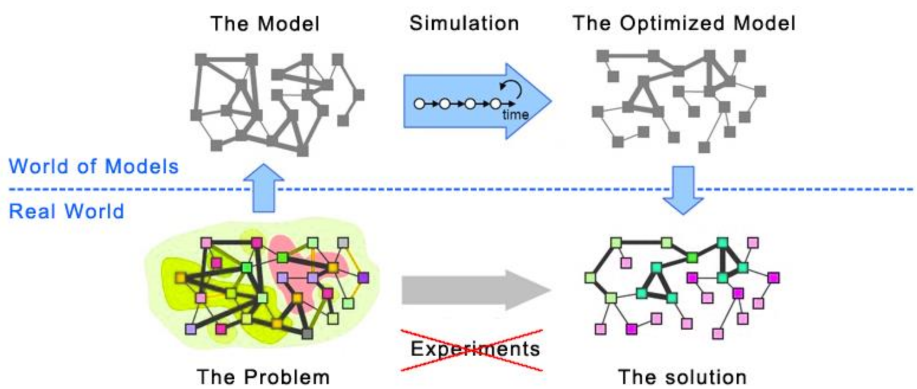

Figure 2. Simulation pipeline. Repeating the actions in real life is expensive or impossible, modeling is used to optimize systems prior to implementation. It includes the stages of mapping the real world problem to the model in the world of models and mapping the solution back to the real system [7].

One of the popular paradigms in simulation technologies is the agent-based modeling & simulation. In agent-based modeling, a system is modeled as a collection of autonomous decision-making entities called agents. Each agent individually assesses its situation and makes decisions on the basis of a set of rules [8]. The agents are autonomous so that they can interact with the environment. The fundamental properties of the agents are the following [9]:

- **Autonomous:** The agent must be able to start or stop an action or restart a stopped action independent from the other agents without any guidance.
- **Reactive:** The agents must be able to observe the environment and act according to the state of the environment.

- **Adaptation:** The agent must be able to plan and manage what actions to take in the changing environment in order to achieve a given goal.
- **Co-operation:** The agent must be able to act together with the other agents in the system, work together with them in order to reach a goal.
- **Temporally continuous**: The current state of the agent should be recorded when it is suspended or terminated. If it is intended to re-animate, it is recreated.

To sum up, the agents are designed in order to serve a purpose. They have control over the internal state of themselves and the actions they take, have the problem solving capability in order to achieve their goal, sense the environment, are the part of the environment. They have well defined interfaces, and are defined clearly [10]. The agents can be implemented by a rule based system or they can also be learning systems [11]. However, since the main requirement of the simulation system is the lack of variety in the data, in the beginning the agents must be rule based systems, though this does not mean that they cannot learn. The system can be improved by constantly observing the real data and changing the agent rules accordingly.

2.3. Data lake

Variety of data is inevitably produced in the city any minute. The only way to make the data valuable is analyzing it. However, gathering the data that is produced by different applications or devices together might not be an easy task since the data can be in different structures. To tackle this task, there should be a system which accepts any kind of data for later processes. Hadoop is an open source project that allows for the distributed processing of large data across clusters of computer which is suitable for solving this problem [12]. It is capable of integrating the sources distributed all around the city and centralizing the data that is perceived by ubiquitous sensors.

Getting the data together is not enough though. The data must be ready for analysis and consumption. While Hadoop is successful at getting the data together, it is not trivial to implement a software that consume the data that sits on a Hadoop cluster [13] [14]. Storing data as itself is not a good choice always. It brings out the need of searching the data interested in the whole dataset, which is costly. For this purpose, one more layer can be added on top of Hadoop. Instant accessibility to the corresponding data for decision support systems is vital in words of managing a city. Traditional analytic systems are lack of the real-time processing on huge amounts of data. These improvements over Hadoop add this real-time capabilities [15].

3. What Makes a City Smart?

In order to raise arguments about how a smart city should be, one needs to understand what makes a city smart. The term smart city is used by many scholar works [16][17][18][19][20].

However since smart city is a relatively new term, there are different usages and there is a need for defining what exactly it means. There are also many other related terms which occur in some works such as Digital City, Intelligent City, Ubiquitous City, Wired City, Hybrid City, Information City, Creative City, Learning City, Humane City, Knowledge City, Smart Community [21]. These terms can be deceiving in a way that one can use the terms interchangeably. However, there are nuances it should be known in order to use these terms correctly. Nam et al. [21] discuss about the differences of these terms in depth and Dameri et al. [22] compares the digital city and the smart city examples. In conclusion, the term smart city can be used for a city if and only if the city meets the criteria for all other terms.

The smartness of the city can be analyzed by defining the application domains of the smart city initiatives. Neirotti et al. systematically define these application domains in [2]. These domains are:

- Energy grids
- Public lighting
- Natural resources
- Water management
- Waste management
- Environment
- Transport, mobility and logistics
- Healthcare
- Public security
- Education and culture
- Social inclusion and welfare
- Public administration and (e-) government
- Economy

These domains can be further divided into subdomains. Although it is possible to approach the emerging problems of getting larger by focusing on these domains separately and invest to the solutions on these domains; it is not necessarily the correct way of solving the problem. This approach may bring further problems in the future as the compatibility issues raise as the number of subsystems increase. Furthermore, these subsystems may require confounding actions to be taken, hence the problem of choosing which action to take arise. Therefore, a centralized system which synchronizes these systems and take care of the orchestration is a must for smart cities.

As another way of elaborating to this requirement, the cities can be analogous to the living organisms. As the veins can be paired with the transportation network, heart can be paired with the city center, and bones can be paired with the city infrastructure components such as the sewage system and the power grids. Besides, the digital ICT infrastructure can be viewed as the nervous system which senses the outside world. In order for body to function, body needs a brain. The brain of the city can be analogous to a centralized system which collects the data from all over the city and either acts by its own or helps the decision makers of the city to make their decisions correctly.

4. Challenges in Integration of Different Domains

As described above, there may be independent components which are responsible for different domains for getting value from the data. All these components have a common goal at the end, which is to increase the people's standard of living. However, since these components are independent of each other, they may distract others and obstruct others' way due to their own mentalities. Making the best decision for the humanity with taking all tradeoffs between these systems into consideration is a real issue because these systems perceive the environment in different ways. An action that result in a positive reward to a component may cause a negative reward to the other.

To illustrate this situation with an example, for a place that is planned to set up a factory, the municipality can refer to the recommendations of smart city components. While one decision support component is specialized for waste management, other can be used for transport and logistics. Their concerns are totally different from each other, where one is in charge of reducing the cost of waste management, the other's task is to reduce the expenses of transportation and logistics. These different viewpoints can turn out as divergent solutions for the same problem. By reason of the fact that this factory is to be set up on a place, it should be the better one. However, only a trusted authority can give the decision which is better. This trusted authority should have an access to the components and the perceptions from the environment to be a wise person in the community.

5. Proposed Solution

As a phenomenon, "smart city" defines a dynamic environment like the human body. Accordingly, the first step to make a city smart is making it digital. Equipping the whole city with cables is not necessary, however, there should be enough perceivers for required places. These perceivers are like the sense organs of the city. The ICT sensors placed all around the city should stream the data they perceived. These perceptions can be in any form from images to audio. They are stored in a centralized data lake for further analyses. This data center can be taught as the collective memory of the city. The stored data is used for learning the pattern that is hidden in everyday life. Since it must be available to integrate the data from various sources, Hadoop is suitable for the implementation of this data lake. For meeting real-time analysis requirements, one more layer can be placed on top of Hadoop as described in the preliminaries section.

Even the city is equipped with the ICT sensors, it needs an intelligent decision-maker mechanism to interpret the sense. The host system is a self-learning system which uses the data to train itself. The learning phase starts from getting relationships out of the events that ICT sensors perceived that the previous decisions caused. This analysis is to just verify whether the decision made is in place or not. While the decision solves the problem of a domain, may arise varied problems in other domains. The calculation of the tradeoff, between these decisions of different domains' components, is learned by host system with reinforcement learning [3]. Their results are visible to the real world and logged in the data center. It is trivial for the host system to extract the reward from the consequences. However, this data may not be enough for considering every possibility. Moreover, repeating the similar actions with different configurations is very costly and not possible most of the time [7]. This situation brings us to the requirement of a realistic agent-based modeling & simulation.

An agent-based simulation is indispensable for the host system to be autonomous. It is substantial for the host system to know how to behave in the situation that has never happened before. Consequently, the host system is in need of imagination ability to act without the guidance of something else. This necessity is satisfied by the agent-based simulation. The simulator gets the decisions from host system as input and generates data like ICT sensors obtained. Host system generates scenarios and sorts them according to their possibilities to emerge. It gets suggestion from subsystems and plays the scenario on the simulator based on each suggestion. In the present instance, environment is not the real city but the city in the simulation. It extracts rewards in compliance with the results of the simulation. Hereby, the host system gains the talent of imagination and the ability of train itself without any help of external sources.

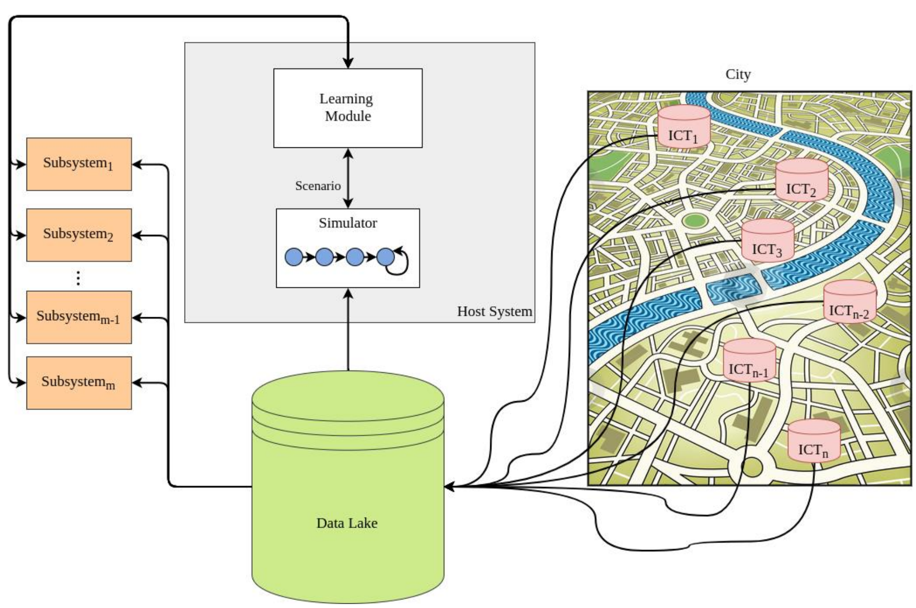

Figure 3. The overall picture of the system. Instead of consuming data directly from the sensors, subsystems use it through the data lake. Simulator and the learning module in the host system operate on the data to figure out the behavior pattern.

The reward used in reinforcement learning for the events that actually happened was obvious. However, it is not so clear for the imagination. The more realistic simulator ends up with more successful host system. Simulation should model the real world close to the truth. This is possible only by imitating the real world. Reinforcement learning can also be applied for training the components of the simulator. The environment is the city with all data that is stored in the data lake. Agents are the main entities of the community such as people, vehicles, companies, roads etc. By replaying the real life again and again in the simulation, these agents can be trained. Their rewards are determined with respect to their consistency to the actual behaviors. In this wise, simulation can be brought closer to the reality.

The tradeoff between the exploration and the exploitation comes to light again. Just like the popular k-bandit problem [4], should the host system try all possible actions or choose the one that it knows best? As explained in the preliminaries section, if the game lasts longer, trying as many choice as possible is a better option. On the development of the autonomous city, training is a lifetime task. Therefore, the host system should prefer the exploration to train itself in its idle time by simulating the city. When somebody ask its advice, it is time to choose exploitation as a short-time strategy. With this workflow, host system continues to develop itself and also serves as its' main purpose.

In the case of adding new components to the system, host system analyzes it and gives priority to it in the simulation to diagnose it. Starting from the most likely scenarios, host system plays the simulation, takes the advices of the new component and observes the consequences of its decisions in various domains. When the host system is satisfied and thinks it is enough to predict the results, it takes the new system's suggestions into account. This pipeline prevents the system wide failures and irreversible faults on the city management.

Conclusions

The gradual increase in urban population brings some problems that are not trivial to solve. Thanks to the advances in the technology, the decision makers of the cities are able to integrate new technologies into their solutions. However, these solutions may be in different domains. Integrating separate technological systems to different domains may not work in practice as they may require taking conflicting actions. In order to make a city not only 'smart' but also 'autonomous', there should be a 'brain' of the city where all the data from the ICT devices is streamed in and the imagination ability for orchestration of the subsystems on various domains is present. The host system runs advanced simulations in order to 'imagine' the results of the events that have never happened before. In this paper, the architecture design of such system is discussed and possible contemporary technologies are compiled. Using the proposed approach, it is probable that the proposed modules can be replaced by more novel innovations that have not yet been discovered.

References

[1] United Nations. (2014). World Urbanization Prospects: The 2014 Revision, Highlights. Department of Economic and Social Affairs. Population Division, United Nations.
[2] Neirotti, P., De Marco, A., Cagliano, A. C., Mangano, G., & Scorrano, F. (2014). Current trends in Smart City initiatives: Some stylised facts. Cities, 38, 25-36.
[3] Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. Journal of artificial intelligence research, 4, 237-285.
[4] Berry, D. A., & Fristedt, B. (2013). Bandit problems: sequential allocation of experiments. Springer Science & Business Media.
[5]Ben-David, S., Blitzer, J., Crammer, K., Kulesza, A., Pereira, F., & Vaughan, J. W. (2010). A theory of learning from different domains. Machine learning, 79(1), 151-175.
[6]C. Voigt and J. Bright, “The lightweight smart city and biases in repurposed big data,” in The 2nd international conference on human and social analytics (HUSO 16), 2016
[7]Borshchev, A., & Filippov, A. (2004, July). From system dynamics and discrete event to practical agent based modeling: reasons, techniques, tools. In Proceedings of the 22nd international conference of the system dynamics society (Vol. 22). Chicago
[8] Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. Proceedings of the National Academy of Sciences, 99(suppl 3), 7280-7287.
[9]Ingham, J. (1997). What is an Agent. Centre for Software Maintenance University of Durham.
[10]Gandon, F. (2002). Distributed Artificial Intelligence and Knowledge Management: ontologies and multi-agent systems for a corporate semantic web (Doctoral dissertation, Université Nice Sophia Antipolis
[11] Van der Hoog, S. (2016). Deep Learning in Agent-Based Models: A Prospectus.
[12] Shvachko, K., Kuang, H., Radia, S., & Chansler, R. (2010, May). The hadoop distributed file system. In Mass storage systems and technologies (MSST), 2010 IEEE 26th symposium on (pp. 1-10). IEEE.
[13] Bittorf, M. K. A. B. V., Bobrovytsky, T., Erickson, C. C. A. C. J., Hecht, M. G. D., Kuff, M. J. I. J. L., Leblang, D. K. A., ... & Yoder, M. M. (2015). Impala: A modern, open-source SQL engine for Hadoop. In Proceedings of the 7th Biennial Conference on Innovative Data Systems Research.
[14] Thusoo, A., Sarma, J. S., Jain, N., Shao, Z., Chakka, P., Zhang, N., ... & Murthy, R. (2010, March). Hive-a petabyte scale data warehouse using hadoop. In Data Engineering (ICDE), 2010 IEEE 26th International Conference on (pp. 996-1005). IEEE.
[15] Liu, X., Iftikhar, N., & Xie, X. (2014, July). Survey of real-time processing systems for big data. In Proceedings of the 18th International Database Engineering & Applications Symposium (pp. 356-361). ACM.
[16] Al-Hader, M., Rodzi, A., Sharif, A. R., & Ahmad, N. (2009, September). Smart city components architicture. In Computational Intelligence, Modelling and Simulation, 2009. CSSim'09. International Conference on (pp. 93-97). IEEE.

[17] Hall, P. (1988). Cities of tomorrow. Blackwell Publishers.
[18] Hall, R. E., Bowerman, B., Braverman, J., Taylor, J., Todosow, H., & Von Wimmersperg, U. (2000). The vision of a smart city (No. BNL--67902; 04042). Brookhaven National Lab., Upton, NY (US).
[19] Mahizhnan, A. (1999). Smart cities: the Singapore case. Cities, 16(1), 13-18.
[20] Partridge, H. L. (2004). Developing a human perspective to the digital divide in the'smart city'.
[21] Nam, T., & Pardo, T. A. (2011, June). Conceptualizing smart city with dimensions of technology, people, and institutions. In Proceedings of the 12th annual international digital government research conference: digital government innovation in challenging times (pp. 282-291). ACM.
[22] Dameri, R. P. (2014). Comparing Smart and Digital City: Initiatives and Strategies in Amsterdam and Genoa. Are They Digital and/or Smart?. In Smart city (pp. 45-88). Springer International Publishing.