Ubinet, Distributed Optimization and Games 2015-2016	December 9, 2015
Lecture 1	
Lecturer: Giovanni Neglia	*Scribe: Giovanni Neglia*

NOTE: The content of these notes has not been formally reviewed by the lecturer. It is recommended that they are read critically.

1 Introduction

In this lesson we introduce 3 levels of abstraction in the description of a distributed system with many agents. The *microscopic* level is concerned with the behaviour of a single agent. The behaviour of a large number of agents can often be described in a simple way in terms of some average quantities. We call this the *macroscopic* description of the system. Finally, it can appear that agents' interactions lead to maximize or minimize some global quantity, this is the *teleological* description of the system.

This classification is proposed in [1]. We illustrate it through an example from Sec. 4.1 of the same reference. We start from an optimization problem on a graph in Sec. 2, showing how it leads to some specific relations among the variables equivalent to Ohm's law for resistor networks. Motivated by this analogy, in Sec. 3 we study a simple model for electrons' movement for which we show Ohm's law describes the aggregate behaviour of a large number of electrons flowing through a resistor. We then summarize our findings in the Sec. 4.

2 A network flow problem

We consider a connected weighted undirected graph. Let $V = \{0, 1, 2, \dots n\}$ denote the set of nodes and E the set of edges. To each edge $(i, j) \in E$ it is associated a non-negative weight $R_{ij} = R_{ji} \geq 0$. We also define $R_{hk} = 0$ if $(h, k) \notin E$. Consider a flow of size U to route from node 1 to node 0. Let I_{ij} denote the flow on link $(i, j) \in E$ and $I_{hk} = 0$ if $(h, k) \notin E$. $I_{ij} \geq 0$ if the flow goes from i to j, $I_{ij} \leq 0$ if the flow goes from j to i. There is a cost $R_{ij} I_{ij}^2$ associated to link (i, j). Our problem is to determine how the flow can be routed from 1 to 0 with the minimum global cost, i.e.:

$$\begin{aligned} \underset{\{I_{ij}\}}{\text{minimize}} \quad & \sum_{i,j} R_{ij} I_{ij}^2 \\ \text{subject to} \quad & \sum_{j \in V} I_{ij} = \begin{cases} U & \text{if } i = 1 \\ -U & \text{if } i = 0 \\ 0 & \text{otherwise} \end{cases} \\ & I_{ij} = -I_{ji} \end{aligned} \tag{1}$$

where the first set of constraints expresses the flow balance at each node, and in particular the fact that 1 is the source and 0 is the sink.

We observe that the function to minimize is convex and all the constraints are linear,[1] it follows that $\mathbf{I}^* \in \mathbb{R}^{|E|}$ is a (global) point of minimum if and only if it is feasible, and there exists constants λ_i^*, $i = 0, 1 \dots n$ and λ_{ij}^*, $i = 0, 1, \dots n$ and $j = i + 1, \dots n$ such that

$$\nabla_{\mathbf{I}} L(\mathbf{I}^*, \boldsymbol{\lambda}^*) = 0 \tag{2}$$

where $\boldsymbol{\lambda}^*$ is a vector containing the constants λ_i^* and λ_{ij}^* and $L(\mathbf{I}, \boldsymbol{\lambda})$ is the Lagrangian

$$L(\mathbf{I}, \boldsymbol{\lambda}) = \sum_{ij} R_{ij} I_{ij}^2 + \lambda_1 \left(\sum_j I_{1j} - U \right) + \lambda_0 \left(\sum_j I_{0j} + U \right) + \sum_{i \neq 0,1} \lambda_i \sum_j I_{ij} + \sum_{i<j} \lambda_{ij} \left(I_{ij} + I_{ji} \right).$$

[1] In this document we use the term *linear* relation also to denote an affine one.

Furthermore, observe that the constraints define a compact set (a polytope) and the cost function is continuous, then the minimization problem has necessarily at least a point of minimum. The point of minimum can then be obtained by solving the system of equations (2). By developing the partial derivatives we obtain:

$$0 = \frac{\partial L}{\partial I_{hk}} = \begin{cases} 2R_{hk}I^*_{hk} + \lambda^*_h + \lambda^*_{hk} & \text{if } h < k, \\ 2R_{hk}I^*_{hk} + \lambda^*_h + \lambda^*_{kh} & \text{if } h > k. \end{cases}$$

If we consider a pair (i, j) with $i < j$, it holds

$$\begin{cases} \frac{\partial L}{\partial I_{ij}} = 2R_{ij}I^*_{ij} + \lambda^*_i + \lambda^*_{ij} = 0 \\ \frac{\partial L}{\partial I_{ji}} = 2R_{ij}I_{ji} + \lambda^*_j + \lambda^*_{ij} = 0 \\ I^*_{ij} = -I^*_{ji} \end{cases}$$

and solving the previous system we obtain that at the coordinates of the point of minimum satisfy the following equations:

$$I^*_{ij} = \frac{\lambda^*_j - \lambda^*_i}{4R_{ij}}.$$

If we define $V^*_i \triangleq -\frac{\lambda^*_i}{4}$, the equation above can be interpreted as the Ohm's law for a resistor with value R_{ij}:

$$I^*_{ij} = \frac{V^*_i - V^*_j}{R_{ij}},$$

and then the optimal solution of problem (1) can be solved by determining the current in a corresponding resistive circuit where a current generator is connected between nodes 0 and 1.

If we revert the point of view we can conclude that in a resistive circuit with an external current generator, the currents are globally solving problem (1) or equivalently, they are minimizing $\sum_{i \neq j} \frac{1}{2} R_{ij} I^2_{ij}$, that is the total energy lost by Joule effect. This is known as Thomson's principle. The currents obey then to *macroscopic* rules like the flow balance at each node and the Ohm's law at each resistor, but at the same time the *teleological* explanation of such laws is that they are moving over the circuits as imposed by the current generator minimizing the global energy loss.

3 A simple model for electrons

We now move to consider a model for electrons' movements in a resistive circuit. Remember that currents arise as the net effect of electrons' movements across resistors. The following numerical example shows that electrons in a conductor move very differently from an ordered army.

Example 1 (Electron Speed). *A typical value for the Fermi energy of an electron in a metal is* $E_F \approx 10eV \approx 10 \times 1.6 \times 10^{-19} J$. *The corresponding speed is*

$$v_F = \sqrt{\frac{2E_F}{m_e}} \approx \sqrt{\frac{2 \times 1.6 \times 10^{-18}}{9 \times 10^{-31}}} \approx 1.9 \times 10^6 \frac{m}{s},$$

where m_e *is the mass of an electron.*

The current in a cylindric wire is $I = nev_{avg}S$, *where* n *denotes the density of free electrons,* S *the surface of the base of the cylinder and* v_{avg} *the average speed of the electrons. If we consider a copper wire* ($n \approx 8.5 \times 10^{28}\ m^{-3}$) *with diameter equal to* 1 *mm, even with a huge current of* $1A$, *the average speed of the electrons in the direction of the current is:*

$$v_{avg} = \frac{I}{neS} \approx \frac{1}{8.5 \times 10^{28} \times 1.6 \times 10^{-19} \times \frac{\pi}{4} \times 10^{-6}} \approx 10^{-4} \frac{m}{s}.$$

It follows then that the average speed of the electrons in the direction of the current is negligible in comparison to its instantaneous speed. Electrons are moving very fast in all the directions with only a small average bias in the direction of the current. This image is very different from an ordered army of electrons all moving with speed v_{avg} *in the direction of the current.*

Because the individual speed of an electron is much higher than the average speed imposed by the electromagnetic field in the conductor, it is possible to imagine that the electron is moving in a random way across the circuit, and only the generator imposes a direction to this chaotic movement. In particular our simplified model assumes that

1. circuit nodes can be occupied by an electron or can be empty,
2. an exponential timer with rate $G_{ij} = \frac{1}{R_{ij}}$ is associated to each resistor R_{ij}, when the timer expires the electrons at the nodes (if any) swap their positions,
3. as soon as an electron arrives at node 1 it is immediately removed by the current generator,
4. as soon as node 0 is empty, the current generator pushes an electron in 0.

The definition of the timers' rates captures the fact that a larger resistor opposes more to the electron passing by it. We define $G_{ij} = 0$ if $(i, j) \notin E$.

We now build an equivalent model. Imagine to have a copy of the original graph (V, E) and to place a token at each node. The token has a black and a white side. The token is placed with the black side upwards if the corresponding node in the electric circuit is occupied by an electron. It is placed with the white side upwards if the corresponding node in the electric circuit is empty. Let the tokens move as follows. When a resistor timer expires, the corresponding tokens are swapped. When a token showing its black face arrives at node 1, it is immediately turned over, as well as any token showing its white face and arriving at node 0. It is immediate to check that the token system evolves as the electron system: whenever the token in i has its black face upwards, an electron occupies site i in the circuit and vice versa.

The advantage of considering the tokens is that the movement of a given token is described by a continuous time Random Walk (RW) on the original graph (ignoring the current generator). In particular the rate at which the token moves from node i to node j is G_{ij}. Because the rates between two nodes are identical ($G_{ij} = G_{ji}$), it follows that the RW is a reversible process (simply check that the uniform distribution $\pi_i = \frac{1}{|V|}$ satisfies the detailed balance equations).

We now wonder what is the probability p_i that the token at node i shows its black face upwards. This is the probability that this token visited node 0 more recently than node 1. Because the token movement is reversible, p_i is also the probability that the token in position i will visit node 0 earlier than node 1. Obviously $p_0 = 1$, $p_1 = 0$. For any other node i, we can calculate p_i conditioning on the possible movements of the token. The token will move to node j with probability $\frac{G_{ij}}{\sum_k G_{ik}}$, then it holds:

$$p_i = \sum_j \frac{G_{ij}}{\sum_k G_{ik}} p_j.$$

For our purposes, it is enough to retain the equation:

$$p_i \sum_k G_{ik} = \sum_j G_{ij} p_j, \text{ for } i \neq 0, 1. \tag{3}$$

We now quantify the net flow of electron across resistor R_{ij}. Observe that p_i is also the probability that there is an electron at node i.

$$\begin{aligned} I_{ij} &= c\left(\Pr(\text{an electron is in } j)G_{ji} - \Pr(\text{an electron is in } i)G_{ij}\right) \\ &= c(p_j - p_i)G_{ij}, \end{aligned}$$

where c is an apposite constant. By defining $V_i = -p_i$, we have found that the currents arising from this simple model for electrons' movement satisfy the Ohm's law. Moreover, the sum of the current at any node i different

It can be checked that current balance at each node is also satisfied. Indeed, for $i \neq 0, 1$

$$\sum_j I_{ij} = c \sum_j p_j G_{ij} - c p_i \sum_j G_{ij} = 0,$$

where the last equality follows from (3). At node 1 the sum of the incoming currents is equal to $-c\sum_j p_j G_{1j}$ or equivalently it is proportional to the rate at which black tokens become white. Similarly, at node 0 the sum of the incoming current is equal to $c\sum_j (1-p_j)G_{1j}$ or equivalently it is proportional to the rate at which white tokens become black. The rate at which a given token changes from white to black is equal to the rate at which it reverts its color from black to white, it follows that the two rates are equal, i.e. $c\sum_j p_j G_{1j} = c\sum_j (1-p_j)G_{1j}$, and the constant c can be determined imposing that these two currents are equal to U.

We have then that the proposed *microscopic* model proposed for electrons' movement in a resistive circuit generates the usual macroscopic equations for currents and potentials.

4 Summary

The example developed during this lesson is an archetype for our course. We have identified three possible levels of description for the considered system. At the microscopic level, electrons are supposed to move according to a simple random walk that takes into account only the local resistors' configuration and the effect of a possible local generator. The aggregation of the movements of many electrons (the macroscopic view) can be described simply in terms of currents' balances at nodes and Ohm's law at each resistor. Moreover, it appears that the electrons, even if unaware, are collectively solving the problem (1), minimizing the loss of energy by Joule effect (the teleological level). This is then the first example of how the local behaviours of many agents can solve a global optimization problem.

The different description levels identified can be helpful to answer different questions about the system. For example the macroscopic description leads to a linear system that is the most effective way to solve the optimization problem (1). At the same time the teleological interpretation allows to answer immediately qualitative questions as the following one: does the circuit dissipate more or less energy if a resistor is added to the circuit? By introducing a new resistor, the solution space is enlarged (there is a new current variable), and then a better solution can be achieved. The total energy dissipated is then smaller. The answer is not evident at all if we restrain to the macroscopic or microscopic descriptions of the system.

References

[1] Frank Kelly and Elena Yudovina, Stochastic Networks. *Cambridge Press*, 2014.