W systemach ICT (ang. *information and communication technology*) rozpowszechnia się *mikroekonomiczny paradygmat projektowania*, wedle którego hierarchiczne architektury zarządzania ustępują miejsca współpracy sfederowanych elementów systemu opartej o indywidualne korzyści. Wśród licznych zwiastunów można wymienić aplikacje rozproszone wspierające pracę grupową, systemy dystrybucji treści multimedialnych, współdzielenie licencjonowanych pasm elektromagnetycznych w systemach LTE, internetowe portale społecznościowe i transakcyjne, systemy samoorganizujące się oraz nowatorskie architektury szeroko rozumianego Internetu Przyszłości. Abstrakcją takich systemów są społeczności inteligentnych *agentów*, które w trakcie interakcji między sobą wymieniają pewne mierzalne dobro zwane *usługami*. Przykładami agentów są operatorzy zasobów, węzły komutacyjne, terminale użytkowników, inteligentne urządzenia, centra danych, roboty softwarowe itp.; przykładowe usługi to udostępnianie zasobów, wyszukiwanie informacji, uczestnictwo lub pośrednictwo w wykonywaniu zadań itp. Fundamentalnym wyzwaniem staje się sprzeczność pomiędzy sposobem działania agentów a ich oczekiwaniami. Z jednej strony agenty są w rosnącym stopniu *autonomiczne* (działają bez nadzoru) i *strategiczne* (maksymalizują własne korzyści antycypując podobne postępowanie innych agentów), ponadto chronią swą prywatność lub nawet anonimowość. Z uwagi na ich dużą liczbę i ochronę prywatności mają niewielką wiedzę o agentach, z którymi wchodzą w interakcje. Sytuacja ta zdaje się wykluczać jakąkolwiek wymianę usług. Z drugiej strony agenty są coraz bardziej od siebie nawzajem zależne, wykonując złożone zadania i wymieniając wrażliwe dane. W społecznościach ludzkich katalizatorem wymiany usług w takich warunkach jest *zaufanie*, formalnie definiowane jako stopień gotowości agenta do uczestnictwa w potencjalnie ryzykownej interakcji. Agenty, które ufają innym, są bardziej gotowe do współdziałania (tj. świadczenia usług), przez co same wzbudzają większe zaufanie – w ten sposób zaufanie sprzyja *uczciwym* zachowaniom. Elementem składowym zaufania jest *reputacja* – ocena zdolności i chęci współpracy agenta, jaką kreuje on w innych agentach poprzez swoje przeszłe zachowania.

Wieloagentowe systemy ICT także zainteresowane są promocją zachowań uczciwych, wypracowując cyfrowe odpowiedniki reputacji i zaufania. Niezbędny do tego jest *podsystem budowy reputacji i zaufania* (ang. *reputation and trust building scheme*, RTBS). Obszary zastosowań RTBS są niezwykle szerokie: od interaktywnych środowisk pracy grupowej i handlu elektronicznego poprzez systemy teleinformatyczne, takie jak radio kognitywne i sieci zdefiniowane programowo, aż po elektroniczne sprawowanie władzy. Znanymi przykładami RTBS są np. system CONFIDANT dla mobilnych sieci *ad hoc*, fora wymiany opinii w portalach Amazon i eBay i in.; do tej kategorii można też zaliczyć mechanizm PageRank firmy Google służący do pozycjonowania stron WWW. Pomimo znacznej obfitości i zróżnicowania istniejących rozwiązań metody systematycznego projektowania RTBS znajdują się ciągle we wczesnej fazie rozwoju. Dlaczego w ogóle są przedmiotem poważnych badań naukowych? Problem polega na tym, że jeśli RTBS mają służyć zastosowaniom komercjalnym, a także krytycznym dla bezpieczeństwa systemów, należy jasno określić, jak bardzo są one podatne na rozmaite manipulacje oraz jak efektywnie realizować ich główny cel, którym jest promocja uczciwych zachowań agentów.

W typowym modelu RTBS agent okresowo wybiera dostawców usług, a następnie raportuje wolumen usług odebranych w trakcie interakcji (tzw. *dane reputacyjne*) do *bloku agregacji reputacji* (ang. *Reputation Aggregation Engine*, RAE), który wylicza i rozpowszechnia *miary zaufania* agentów dla wspomagania ich przyszłych decyzji o świadczeniu i raportowaniu usług. *Nieuczciwy* agent może zaatakować RTBS z pobudek egoistycznych lub złośliwych, np. dla sztucznego podnoszenia swej miary zaufania przy świadczeniu niewielu usług. Może też sztucznie podnosić wiarygodność własnych, bądź obniżać wiarygodność cudzych danych reputacyjnych. Niektóre ataki stają się bardziej niebezpieczne w warunkach *zmowy* agentów nieuczciwych. Obecnie poszukuje się metod systematycznego projektowania RTBS i ich uodporniania na nieuczciwe zachowania agentów strategicznych. Można to uznać za zagadnienie z zakresu "miękkich" środków bezpieczeństwa (ang. *soft security*), gdzie tradycyjne "twarde" środki blokujące dostęp nieuczciwych agentów do systemu zastępuje się zniechęcaniem do zachowań nieuczciwych bądź ograniczaniem ich skutków.

Trudności projektowania RTBS wzrastają, gdy wybór dostawcy usług jest "ślepy" – dyktowany bliskością geograficzną lub dostępnością usług, nie zaś miarą zaufania. Ta częsta sytuacja rzadko jest uwzględniana i wymaga nowych metod projektowania. Jedyną formą promocji zachowań uczciwych jest wówczas różnicowanie polityki świadczenia i raportowania usług w zależności od miary zaufania partnera w interakcji, tj. nagradzanie zachowań uczciwych w stosunku do agentów trzecich zgodnie ze znaną w biologii, ekonomii i socjologii *zasadą wzajemności pośredniej*. Analityczne i symulacyjne badania z wykorzystaniem metod interdyscyplinarnych w stosunku do systemów ICT w warunkach anonimowości agentów i "ślepego" wyboru dostawców usług powinny odpowiedzieć na szereg fundamentalnych pytań, takich jak: Czy agenty nieuczciwe mogą uzyskiwać wyższe miary zaufania niż agenty uczciwe (jeśli tak, RTBS traci rację bytu, gdyż wysokie miary zaufania przestają być pożądane)? Co należy zrobić, by uniknąć takiego scenariusza? Jaki wpływ na to ma liczba agentów nieuczciwych oraz fakt ich zmowy lub braku zmowy? Na jaki wolumen usług odbieranych mogą liczyć agenty uczciwe w obecności egoistycznych bądź złośliwych agentów nieuczciwych? Czy wzajemność pośrednia jest opłacalna? Czy warto utajniać algorytm działania RAE?